

« LA LEXICOTHÈQUE »

— o O o —

LA MARQUE EN LEXICOGRAPHIE

États présents, voies d'avenir

Textes réunis et présentés par Fabienne Baidier,
Efi Lamprou et Monique Monville-Burston



Lambert-Lucas
L I M O G E S

La marque lexicographique concerne aussi bien l'aspect sociolinguistique que l'aspect sémantique des mots. Comme le suggère le présent recueil, fruit du colloque international qui a réuni des spécialistes du français et du grec, mais aussi d'autres langues (anglais, tchèque, kaingang, espagnol...), ce double emploi soulève de nombreux débats.

La première partie, consacrée à la sémantique, traite des descriptions lexicographiques et des problèmes d'équivalences dans les dictionnaires de langue, y compris en ce qui concerne la traduction.

La deuxième partie, d'ordre sociolinguistique, examine le marquage diastratique réalisé par *arg.*, *fam.* et *pop.* et le marquage diatopique des variétés belge, chypriote, québécoise et suisse du français. De tels classements géolinguistiques sont-ils fondés ? Quel est le statut de la notion de « variante » ? Plus généralement, le marquage du lexique « non conventionnel » répond-il aux réalités linguistiques actuelles ? Comment les valeurs signalées sont-elles marquées, décrites et distribuées dans un dictionnaire électronique ?

Des propositions méthodologiques basées sur de nouveaux cadres théoriques sont présentées en fin de volume.

Contributions de Dorothée Aquino-Weber, Anna Anastassiadis-Syméonidis, Jean-Claude Anscombre, Haifa Ben Mahfoudh-Hubert, Erzébet Chmelik, Pierre Corbin, Christophe Cusimano, Jean-Nicolas De Surmont, Lola-Laurence Devolder, Mireille Elchacar, Vassiliki Foufi, Nathalie Gasiglia, Tita Kyriakopoulou, Efi Lamprou, Claude Martineau, Camille Martinez, Louis Mercier, Christel Nissile, Pavlos Parlou, Bert Peeters, Alena Podhorná-Polická, Pierre-Yves Raccah, Jackie Schön, Freiderikos Valetopoulos, Anastasia Yiannacopoulou.

270 pages

30 euros

ISBN : 978-2-915806-98-4

Collection « La Lexicothèque » ISSN : 0753-3454

LA MARQUE EN LEXICOGRAPHIE

États présents, voies d'avenir

Textes réunis et présentés
par Fabienne BAIDER, Efi LAMPROU
et Monique MONVILLE-BURSTON

Contributions de
Dorothee ACQUINO-WEBER
Anna ANASTASSIADIS-SYMÉONIDIS
Jean-Claude ANSCOMBRE
Haifa BEN MAHFOUDH-HUBERT
Erzébet CHMELIK
Pierre CORBIN
Christophe CUSIMANO
Jean Nicolas DE SURMONT
Lola-Laurence DEVOLDER
Mireille ELCHACAR
Vassiliki FOUI
Nathalie GASIGLIA
Tita KYRIAKOPOULOU
Efi LAMPROU
Claude MARTINEAU
Camille MARTINEZ
Louis MERCIER
Christel NISSILE
Pavlos PAVLOU
Bert PEETERS
Alena PODHORNÁ-POLICKÁ
Pierre-Yves RACCAH
Jackie SCHÖN
Freiderikos VALETOPOULOS
Anastasia YIANNACOPOULOU

*Ouvrage publié avec le concours
de l'Université de Chypre*



Actes du Colloque international de l'Université de Chypre
Département d'études françaises
Nicosie, 21, 22 et 23 octobre 2006

© Lambert-Lucas, Limoges, 2011

ISBN : 978-2-915806-98-4

Collection « La Lexicothèque » ISSN : 0753-3454

INTRODUCTION

par Fabienne Hélène BAIDER, Efi LAMPROU
et Monique MONVILLE-BURSTON

La marque lexicographique peut être définie comme un signe qui « explicite la valeur d'emploi d'un mot ou d'un terme sur la base de l'analyse des connotations » (Rousseau 2008) ¹. Les deux notions « valeur d'emploi » (aspect sociolinguistique de la marque) et « connotation » (aspect sémantique de la marque) ont soulevé et soulèvent toujours des débats. Notre première partie, consacrée aux ASPECTS SÉMANTIQUES, analyse et met en cause la notion même de marque afin de faire de nouvelles propositions. La deuxième partie, consacrée aux VALEURS D'EMPLOIS, aborde les problèmes plus spécifiques relatifs aux marques topocentrales (Suisse, Belgique, Québec et Chypre) et sociolectales (argotique, populaire, familial, savant).

PREMIÈRE PARTIE ASPECT SÉMANTIQUE

Ce premier volet se divise en deux : *Typologie* fait un état présent des descriptions lexicographiques dans les ouvrages français et grecs ; *Équivalence* est consacrée aux problèmes de traduction de la connotation et de la dénotation dans les ouvrages lexicographiques.

I. TYPOLOGIE

Dans le premier article, Pierre Corbin et Nathalie Gasiglia décrivent et évaluent les progrès effectués dans les modalités du marquage lexicographique afin de cerner l'impact de l'évolution des conceptions et des connaissances sociolinguistiques et terminologiques, entre autres. Deux questions constituent le fil conducteur de ce très riche article. Tout d'abord, sur quels principes le marquage s'appuie-t-il ? Ensuite, quelles pratiques prennent véritablement effet ? Les auteurs prennent pour point de départ les résultats d'une étude antérieure (Corbin 1989), et s'appuyant sur la typologie des onze types de marques *en lexicogra-*

¹ *Les marques terminographiques : une approche aménagiste*, sur le site Realiter, Direction Terminologie et Industries de la Langue - Union latine, Réseau panlatin de terminologie, article 372, www.realiter.net/spip.php?rubrique18

phie générale et monolingue (Hausmann 1989) établissent les progrès effectués. La première étude avait en effet repéré un déficit d'explicitation et de cohérence relatif aux modalités de ce marquage. Travaillant avec le même corpus de dictionnaires, dictionnaires généraux et ouvrages plus spécifiquement destinés aux jeunes apprenants natifs, les auteurs concluent à des progrès dans le domaine : explicitation des principes plus systématique, marquage des variations diaphasiques plus justes et augmentation des francophonismes dans la nomenclature. Ce bilan nécessaire et fort utile invite cependant à « désidéologiser » le marquage diaphasique et à s'attacher de manière plus rigoureuse à appliquer des principes lors de l'assignement de marques.

L'impact des conditions économiques et techniques sur le marquage fait l'objet de l'article de Camille Martinez, qui s'appuie sur une comparaison des *Petit Larousse* millésimés 1997 à 2007 pour décrire le processus de mise à jour de ce dictionnaire. Si tous les marqueurs lexicographiques sont largement touchés lors des mises à jour, l'étude se focalise sur le marqueur *littéraire* dans la lettre E des *Petit Larousse* 1997 à 2007. Cette étude affirme que les apparitions et les disparitions de marqueurs ne sont pas nécessairement motivées par des raisons linguistiques (informations nouvelles, désambiguïsation, etc.). En effet, dans ce corpus, des contraintes d'ordre spatial (et donc économique) jouent aussi un grand rôle puisque la plupart des changements se produisent à proximité d'ajouts d'articles nouveaux. Cependant, cette étude met aussi en lumière que l'aspect humain – plutôt que mécanique – du travail lexicographique est un facteur qui génère de l'incertain dans la motivation du retrait, de la modification ou de l'ajout des marques.

Cette incertitude motive aussi le travail d'Anna Anastassiadis-Syméonidis consacré aux marques utilisées dans le *Dictionnaire du grec standard* ou DGS (Institut d'Études Néohelléniques, 1998) et le *Dictionnaire du grec moderne* ou DGM (Babiniotis 1998). Ce corpus est homogène dans la mesure où les ouvrages sont comparables en plusieurs points (monolingues, publiés en 1998, en gros de la même taille et s'adressant au même public). Après un tour d'horizon fort instructif de différentes typologies proposées, en particulier celle de Rey-Debove (1971) et celle de Coseriu (1973), l'auteure concentre son étude sur les marques d'usage diaphasique et diastratique (savant, officiel, populaire, familial, argotique, vulgaire). La comparaison de vingt entrées dans les deux dictionnaires invite Anna Anastassiadis-Syméonidis à constater que ces marques, relevant des « conditions d'énonciation », constituent une source d'incohérence notable dans le système de marquage. Ayant elle-même participé au marquage dans le DGS, l'auteure émet trois propositions concrètes pour améliorer ce système : (a) procéder à une définition biunivoque des marques ; (b) bannir des dictionnaires les jugements normatifs ; (c) procéder à une description plus détaillée des unités lexicales au niveau du marquage en tenant compte de trois tranches d'âge selon lesquelles nous diviserions la société (jusqu'à 25 ans, jusqu'à 60 ans, au delà).

Les principes méthodologiques relatifs à l'*équivalence* font l'objet des travaux que nous avons rassemblés dans la deuxième section. Cette équivalence sémantique est abordée sous différents angles complémentaires, que ce soit la traduction de proverbes, la problématique des faux-amis et celle de la place des présupposés idéologiques dans les dictionnaires.

II. ÉQUIVALENCE

Jean-Claude Anscombre expose les problèmes auxquels il se trouve confronté en tant que traducteur des formes sentencieuses en vue de l'élaboration d'un dictionnaire bilingue espagnol-français. Après avoir présenté un certain nombre d'exemples de traductions considérées « ratées », l'auteur met au jour les principes implicites à l'origine de la médiocrité des équivalences proposées : (a) ce qui a été autrefois un proverbe est toujours un proverbe ; (b) étant donné un proverbe d'une langue L_1 donnée, tout énoncé d'une langue L_2 qui lui ressemble et qui est un proverbe fournit une équivalence valable du premier ; (c) c'est dans les compilations et dictionnaires antérieurs que l'on trouve des proverbes « qui se ressemblent ». Pour améliorer les propositions d'équivalence, l'auteur présente d'autres principes méthodologiques et les applique à des exemples concrets. Si l'équivalence catégorielle est importante pour l'insertion du paramètre pragmatique, l'auteur souligne que l'équivalence lexicologique, l'équivalence statistique et l'équivalence stylistique (respect du niveau de langue et des variantes régionales) doivent aussi être prises en considération.

Ce problème d'équivalence est aussi abordé par Bert Peeters pour qui, derrière la grande majorité des faux-amis, se cachent des similarités sémantiques qui risquent d'échapper à la vigilance des lexicographes. En effet, après avoir établi que les faux amis étaient des paires de mots de langues différentes et de sens partiellement ou totalement différents, en dépit d'une ressemblance aussi bien du point de vue de l'orthographe que de celui de la prononciation, l'auteur étudie les « faux amis à différents degrés ». Travaillant à l'aide de l'outil descriptif qu'est la métalangue sémantique naturelle d'Anna Wierzbicka, Bert Peeters explicite des différences entre des faux-amis tels que *évidence* et *expérience* (français) et *evidence* et *experience* (anglais) à partir de phrases dites *canoniques*, qui illustrent l'usage qu'on peut faire des éléments lexicaux qui constituent le lexique de la métalangue. L'auteur établit que cette métalangue permet aussi la recherche de similarités entre expressions, et par là même une meilleure compréhension du rapport qui existe entre des faux amis particuliers. En effet, ceux-ci ne sont pas toujours « aussi faux », ni même toujours « aussi partiels » qu'il y paraît. C'est effectivement sur les *différences* que les analyses antérieures ont insisté le plus souvent, alors que le changement de perspective permet d'identifier un noyau sémantique commun.

Christophe Cusimano s'interroge quant à lui sur le bien fondé des traductions de langues peu connues telles que le kaingang, langue amérindienne, et il avance quelques propositions dans le cadre des dictionnaires bilingues. Ainsi aux traductions données par le dictionnaire, il oppose les résultats d'une enquête de terrain réalisée auprès de locuteurs natifs. Il met ainsi en évidence que lorsque deux langues ont un statut différent et que le lexicographe appartient à celle des deux qui jouit d'une meilleure reconnaissance sociale, la langue traduite subit l'effet des habitudes culturelles des traducteurs dans leur choix de traductions. Un exemple, celui d'une linguiste allemande qui s'est appliquée à décrire le kaingang, nourrit cette démonstration.

Enfin Pierre-Yves Raccach et ses collaboratrices, Haïfa Ben Mahfoudh-Hubert et Erzsébet Chmelik, explicitent justement comment prendre en compte dans un dictionnaire de langue les présupposés culturels évoqués dans le travail

de Cusimano et d'Anscombe, présupposés qui sont ici appelés *présupposés idéologiques cristallisés dans le lexique* (PICs). Ces PICs sont des points de vue dont les mots de la langue imposent l'acceptation préalable à la compréhension des discours et des textes qui les utilisent. Ce travail s'inscrit dans un programme de recherche plus large dont le but est de fournir un cadre théorique efficace et fiable pour la construction d'un dictionnaire de langue dans lequel on rendra compte des PICs. Les auteurs démontrent à partir de l'exemple de *jouer*, qu'il est possible de fournir, de manière systématique, des éléments de description des points de vue présupposés dans les mots du lexique, et ce, en utilisant un modèle descriptif de ce que l'on peut appeler une « sémantique instructionnelle ». Cette étude suggère des pistes à explorer pour approfondir les recherches et/ou pour trouver des dispositifs permettant aux descriptions *lexicologiques* et *lexicographiques* du mot d'être complétées de manière à inclure ces contraintes.

Avec ce dernier article s'achèvent les réflexions sur la description du sens proprement dit dans les dictionnaires. Les travaux de la deuxième partie illustrent les problèmes relatifs à la marque diatopique et notamment aux variétés de français québécoise, suisse, belge et chypriote. Ces travaux soulèvent des problèmes de méthodologie et certains illustrent des essais de nouveaux cadres théoriques permettant l'établissement des systèmes des marques d'usage.

DEUXIÈME PARTIE VALEURS D'EMPLOIS

I. MARQUE DIATOPIQUE

Lors de leurs travaux sur le volet suisse romand de la *Base de données lexicographiques panfrancophone* (BDP)² à l'Université de Neuchâtel, Christel Nissile et Dorothée Aquino se sont interrogées sur le statut de l'emploi des lexèmes dits « variantes » et sur le bien-fondé de leur classement géolinguistique. De manière concrète, cet article présente un « outil » qui permettrait d'appréhender la réception et la production de lexèmes régionaux, en l'occurrence ici suisses romands. La première étape consiste en l'établissement du profil de chaque lexème (représentation géographique, fréquence et vitalité ou marque diachronique, diatopique, diafréquentielle) pour les différentes zones linguistiques de la Suisse romande. Dans la deuxième étape il s'agit de tester à la fois l'usage réel des locuteurs/scripteurs – la norme de production – et les représentations conscientes et inconscientes que le locuteur se fait de sa langue (établissement des marques diatextuelle, diamésique, diaphasique et diatechnique). Cette étape devrait permettre la constitution d'une échelle de variation diastratique qui servirait à dégager une « norme » régionale objectivée, et à mettre en lumière, dans l'axe de variation diatopique, l'existence de centres de référence et d'irradiation qui jusqu'à présent n'étaient que pressentis dans les travaux antérieurs. L'étude des lexèmes *roiller*, *aguillage*, *huitante*, *encoubler*, *prétérirer*, entre autres, illustre de manière précise l'emploi de ce nouvel outil.

La fonction identitaire des variétés régionales est aussi au centre des travaux de Mireille Elchacar et de Louis Mercier, chercheurs et experts en lexicographie québécoise au sein du FRANQUS (Université de Sherbrooke). Leur rétrospective

² La BDP consiste en la mise en réseau des différentes descriptions lexicographiques différentielles des variétés régionales du français.

de la présentation et de l'inclusion des particularismes québécois dans la lexicographie québécoise, exemplifiée par les lexèmes *sucre*, *moufle*, *mitaine*, démontre de manière convaincante que cet historique est aussi l'histoire d'une affirmation identitaire grandissante. Après avoir été limitée à des ouvrages normatifs et s'être focalisée sur ce qui la démarque du « français standard », la lexicographie québécoise tâche depuis un quart de siècle de concrétiser un ouvrage général consacré à la description du français québécois standard. Celui-ci doit non seulement répondre aux besoins langagiers des francophones du Québec et du Canada, mais aussi respecter leur expérience et leur point de vue de Nord-Américains. Dans un tel ouvrage il s'agit de marquer les usages européens par rapport à l'usage québécois, et, ainsi, de promouvoir la cohabitation d'usages québécois et français grâce à des marques topolectales. Au moyen d'exemples appartenant au vocabulaire politique est alors étudiée et expliquée la manière dont ces marques topolectales peuvent être utilisées pour mettre en relation des réalités politiques différentes dans un dictionnaire général.

Le français de Belgique fait l'objet d'un intérêt grandissant comme en témoigne l'inclusion de plus en plus importante de belgicisms dans le *Petit Larousse illustré*. Jean Nicolas De Surmont, à partir d'un corpus de belgicisms, étudie les problèmes soulevés par leur marquage lexicographique, leur vitalité, leur sélection et leur traitement lexicographique. L'auteur note en particulier la vitalité relative en chute de certains belgicisms (donc non représentatifs), l'emploi dans une autre variété nationale de français d'un terme marqué « belgicisme » et la haute variabilité d'un dictionnaire à l'autre des pratiques d'inclusions dans la nomenclature et de marquages dans la microstructure. Le corpus se limite aux divergences de traitement et de sélection dans quelques dictionnaires récents de langue française, tels que le *Dictionnaire Hachette*, le *Petit Larousse illustré* et le *Grand Robert électronique* et évoque des variétés souvent ignorées dans les études, celles d'Afrique et d'Océanie.

Pavlos Pavlou consacre sa discussion aux emprunts ou mots marqués dans la mesure où ils sont employés par une certaine génération : l'auteur pose ainsi la question de l'inclusion d'emprunts récents à la langue anglaise dans la nomenclature d'un dictionnaire du chypriote grec, emprunts dont l'usage est limité à une partie d'une communauté (les jeunes). De plus, non seulement le facteur âge joue un rôle déterminant dans la connaissance et l'usage de ces emprunts, mais le facteur sexe fait aussi varier la vitalité de ces emprunts quel que soit l'âge (ils sont plus employés par les hommes). Marquer de tels mots dans un dictionnaire pourrait se révéler complexe et pourrait surtout aller à l'encontre de la philosophie des dictionnaires généraux, conçus pour représenter l'idiome de la majorité de la population, ou du moins la partie de l'idiome stable dans le temps et dans l'espace. Ce dernier article met en avant la combinaison de deux facteurs, celle d'une variété régionale (marque topolectale) et celle d'une génération et d'un sexe (marque sociolectale), ce qui nous amène au dernier volet de notre volume, consacré à la marque sociolectale.

II. MARQUE DIASTRATIQUE

Cette section met en évidence la difficulté de marquer les différents sociolectes. Ainsi les marques dites d'usage, utilisées en lexicographie afin de rendre compte de l'hétérogénéité du lexique, se limitent-elles à *arg.* (argotique), *fam.* (familier)

et *pop.* (populaire). Que ce soit dans une perspective stylistique ou sociolinguistique, ces marques font l'objet de débats quant à leur définition et à leur fonction de différenciation.

De fait, Lola Devolder observe qu'une comparaison diachronique entre les marques établies dans trois grands dictionnaires monolingues français (le *Lexis*, le *Petit Larousse* et le *Petit Robert*) à trente ans d'intervalle ne laisse aucun doute : la répartition des trois marques *arg.* (argotique), *fam.* (familier) et *pop.* (populaire) est encore très disparate d'un dictionnaire à l'autre. Pour exemple, le verbe *jaspiner* sera marqué *fam.* et *péj.* par le *Petit Robert* 2003, *pop.* par le *Lexis* 2002 et *arg.* par le *Petit Larousse* 2003. Le problème se situe-t-il au niveau des concepts ou vient-il des choix différents des comités éditoriaux ? C'est cet imbroglio fait de « bas langage », de « langue verte », de lexique « non conventionnel » ou simplement de « français commun », que l'auteure propose de démêler. Dans un premier temps est examinée la présentation de ces marques dans les préfaces des dictionnaires à l'aide de la distinction diaphasique / diastématique, préfaces trop souvent ignorées des travaux sur le discours lexicographique. Ensuite l'analyse des critères étymologique, morphologique, sémantique et sociolinguistique permet de critiquer l'attribution de chacune des marques sociolectales dans un corpus limité aux entrées sous les lettres J et K. La conclusion est qu'au regard des réalités linguistiques actuelles, la pertinence d'une telle catégorisation peut être remise en cause.

De même, Alena Podhorná souligne la confusion terminologique lors de la définition du niveau de langue qu'est l'argot, que ce soit chez les linguistes français ou les linguistes tchèques, confusion typique venant, selon l'auteure, d'un flou général concernant le classement du lexique non standard. D'une part, les lexicographes n'adoptent pas de critères univoques ; d'autre part, les lexicologues reformulent les critères de classement selon des points de vue variables. La comparaison de deux éditions d'un même dictionnaire montre que les marques « vulgaire », « argotique », « populaire » et « familier » ne fonctionnent pas comme des catégories indépendantes. De plus, le social et le situationnel jouant un rôle notable dans la dynamique de l'usage des expressions appartenant à ces registres, il en résulte une grande difficulté pour analyser de tels lexèmes en temps réel. Deux solutions sont alors proposées : soit l'adoption par tous les dictionnaires de la nouvelle marque combinée, *arg. fam.* du *Petit Robert* ; soit l'insertion de la marque « expressive », employée dans la lexicographie tchèque, qui permettrait de contourner un bon nombre de problèmes liés à la catégorisation des cas limites.

Le travail de Jackie Schön, consacré aux acceptions familières désignant en particulier des êtres humains, prend aussi en compte la situation de communication pour établir la marque lexicographique. Ainsi l'auteure propose-t-elle d'associer la marque *FAM.* à l'expression de la subjectivité des locuteurs à l'égard des personnes ainsi désignées. En effet, une acception apparaît marquée si elle désigne – pour une forme donnée – un nouveau sens qui s'ajoute à son sens premier. L'auteur considère notamment les lexèmes *andouille* ou *cruche* dont les acceptions secondes (ou « marquées ») sont respectivement illustrées par des séquences telles que *Sacrée andouille ! Quelle cruche !* Le sens attribué à ces acceptions devient communément : « niais », « imbécile ». Le mécanisme à l'œuvre dans ce transfert de sens et de domaine produit la familiarité qui naît

donc d'une incongruité. Cette incongruité provient du recouvrement systématique d'une classe de dénominatifs des êtres par des classes sémantiques de noms impropres à leur dénomination (noms de comestibles, d'objets domestiques, etc.). La familiarité langagière est avant tout un phénomène discursif, mais suffisamment codé et stable pour être répertorié dans les dictionnaires de langue sous les traits de la marque et c'est justement ce que propose ce travail.

Les quatre marques « familier », « argotique », « populaire » et « vulgaire » font aussi l'objet des recherches d'Efi Lamprou et de Freiderikos Valetopoulos. Leur étude des dictionnaires bilingues bidirectionnels français / grec part de l'interrogation suivante : Qu'attendent les usagers d'un dictionnaire bilingue en ce qui concerne les niveaux de langue non standard ou non conventionnels ? Les dictionnaires répondent-ils à ces attentes ? Tout d'abord à partir de la description précise des microstructures, les auteurs s'interrogent sur la pertinence de certaines informations présentées dans les ouvrages lexicographiques, telles que les marques assignées, ainsi que sur l'exactitude des équivalents proposés. Tous les exemples sont tirés de dictionnaires français/grec et cette étude a l'originalité de recouvrir deux champs de recherche de notre volume : celui de l'équivalence de la connotation dans la traduction, ainsi que celui du bien-fondé de l'emploi des marques dans des corpus unilingues.

Des corpus de langue française et de langue grecque font aussi partie de l'analyse proposée par l'équipe du Laboratoire d'Informatique de l'Institut Gaspard-Monge (IGM). Celle-ci étudie l'utilité de marques sociolectales et stylistiques dans le cadre d'un dictionnaire électronique, mais aussi les difficultés à les coder. Pour ce faire, les auteurs s'appuient sur deux cas concrets puisés dans le Dictionnaire électronique du français et le Dictionnaire électronique du grec moderne en élaboration à l'Université de Thessalonique. Dans le cas du dictionnaire français, la problématique est illustrée par des exemples appartenant à trois couches lexicographiques (mots courants, mots moins courants et mots rares ou techniques). En ce qui concerne la langue grecque, la marque lexicographique qui leur sert d'exemple est la marque +sv (« forme savante ») trouvée dans le dictionnaire grec. En effet, la langue grecque est marquée par la diglossie entre la langue savante (katharevoussa) et la langue démotique, ce qui complique à la fois la tâche des linguistes et des informaticiens, une difficulté à laquelle cette équipe cherche à trouver une solution.

Connotation, emploi, marque ne sont jamais, semble-t-il, une question close. C'était de fait la raison d'être de notre rencontre internationale en octobre 2006 mais cela en est aussi la conclusion. En effet, les travaux consignés dans ce volume font plusieurs propositions relatives au marquage et à l'équivalence connotative et dénotative. Celles-ci devraient nous permettre de dépasser par exemple les idées traditionnelles véhiculées par des notions comme « régionalisme » ou « sociolecte » pour considérer que les registres dits « marqués » s'inscrivent dans un diasystème complet plutôt qu'en marge de la langue.

PREMIÈRE PARTIE
ASPECT SÉMANTIQUE

I.
TYPOLOGIE

ÉLÉMENTS POUR UN ÉTAT DE LA DESCRIPTION DE LA VARIÉTÉ DES USAGES LEXICAUX DANS LES DICTIONNAIRES FRANÇAIS MONOLINGUES (1980-2008)

Pierre CORBIN et Nathalie GASIGLIA ¹
Université Charles de Gaulle - Lille 3
UMR 8163 du CNRS, « Savoirs, Textes, Langage »

Résumé — Cet article donne une vue générale des différents principes de marquage appliqués depuis le milieu des années 1980 dans les dictionnaires de langue et d'apprentissage du français. Pour autant que les dictionnaires de langue soient concernés – les premiers d'entre eux ont été décrits antérieurement –, on repère plusieurs types d'évolutions sur un fond de pratiques stables, sinon statiques. En ce qui concerne les dictionnaires d'apprentissage – un concept relativement récent –, nous proposons une description initiale des moyens qu'ils emploient pour caractériser les divers usages lexicaux pris en compte.

Mots clés — dictionnaires généraux, dictionnaires d'apprentissage, variation, marques, marquage.

Abstract — Our article offers an overview of various principles of labelling implemented from the mid-eighties in general and learner's French dictionaries. As far as general dictionaries are concerned, already described for the earlier period, we point out several evolution processes against the background of stable, if not static, practices. For learner's dictionaries, mainly of rather recent conception, we propose an initial description of the means they use to characterize a variety of lexical usages taken into account.

Keywords — general purpose dictionaries, learner's dictionaries, language variety, labels, diasystematic labelling.

¹ Nous remercions Jean Pruvost pour son aide documentaire, Katia Paykin-Arrouès pour la traduction du résumé en anglais, ainsi que les évaluateurs anonymes et les organisatrices du colloque pour leurs remarques pertinentes, que nous avons mises à profit.

INTRODUCTION

Au traitement du sens et de la combinatoire des items lexicaux qu'ils répertorient, les dictionnaires associent un appareil empirique de description de la distribution de leurs usages dans les pratiques des locuteurs et d'évaluation de ceux-ci selon divers paramètres. La partie la plus visible de cet appareil est un jeu de codifications textuelles et typographiques appelées « marques d'usage », mais il peut aussi, selon les besoins et les circonstances, prendre la forme de notations discursives diversement développées intégrées aux définitions ou constituant la matière de remarques.

À la suite des études analytique de D. & P. Corbin (1980) et synthétique de P. Corbin (1989), qui ont donné une vue des positionnements et des pratiques dictionnaires français des années 1960 au milieu des années 1980, nous avons engagé, dans le cadre du master professionnalisant « Lexicographie, terminographie et traitement automatique des corpus » que nous animons à l'université Lille 3², une recherche extensive sur les constantes et les évolutions observables dans les deux dernières décennies, dont nous présentons ici quelques aperçus.

Nous utiliserons comme cadre pour les développements qui suivent la typologie de l'appareil de description dictionnaire des usages de Hausmann (1977, chap. 8) enrichi par Hausmann (1989, § 4), qui distingue onze sortes de marquage articulées sur autant de critères :

- *diachronique* : temporalité régressive (pour les usages vieillissants ; ex. : *vx*, *vieilli*) ou progressive (pour ceux qui apparaissent ; ex. : *néol.*) ;
- *diatopique* : spatialité intraterritoriale (pour les usages hexagonaux ; ex. : *région.*, *dial.*) ou extraterritoriale (pour ceux de la francophonie ; ex. : *Belgique*, *Suisse*, *Canada*) ;
- *dia-intégratif* : idiome d'origine des emprunts (en tant qu'il fait l'objet de considérations extérieures à la rubrique étymologique des articles ; ex. : *anglicisme*) ;
- *diamédial* : médialité (« canal » d'expression oral ou écrit ; ex. : *oral*) ;
- *diastratique* : groupe socioculturel auquel appartiennent les locuteurs (ex. : *pop.*) ;
- *diaphasique* : formalité plus ou moins grande des usages (ex. : *fam.*, *soutenu*) ;
- *diatextuel* : type de discours ou genre textuel (ex. : *admin.*, *didact.*, *littér.*, *poét.*, *sc.*) ;
- *diatechnique* : technicité (toutes les marques de domaine ; ex. : *chim.*, *dr.*, *méd.*, etc.) ;
- *diafréquentiel* : fréquence (ex. : *rare*) ;
- *diaévaluatif* : attitude énonciative des locuteurs (ex. : *euphém.*, *iron.*) ;
- *dianormatif* : normativité (ex. : *emploi critiqué*).

En dépit de discussions diverses auxquelles pourrait donner lieu ce classement (concernant le découpage opéré, les choix dénominatifs effectués, le manque de hiérarchie et d'articulation entre les critères retenus ou encore la distribution des marques sur les différents types³) mais qui ne font pas l'objet de la présente contribution, il sera retenu comme repère pour les développements qui suivent en vertu de sa relative notoriété, dont témoigne la diffusion de sa terminologie⁴, et de son utilité pour la comparaison des répertoires en l'absence d'un modèle sociolinguistique alternatif convaincant pour le français⁵.

Dans ces quelques pages, qui ne permettent pas d'entrer dans le détail des pratiques actuelles de marquage ni de couvrir l'ensemble des dictionnaires français (les répertoires spécialisés seront laissés de côté), nous nous limiterons à dégager certaines lignes de force, en nous attachant en priorité aux principes, explicites ou implicites, qui sont mis en œuvre dans les dictionnaires généraux pour utilisateurs francophones relativement experts et les ouvrages plus spécifiquement destinés aux allophones ou aux jeunes apprenants natifs.

1. DICTIONNAIRES GÉNÉRAUX POUR LOCUTEURS CONFIRMÉS

Par rapport à la période qui va jusqu'au milieu des années 1980, l'évolution récente des dictionnaires généraux monolingues français est marquée par des progrès dans l'explicitation des principes de description de la variété des usages, par une régression relative du marquage diastatique au profit du diaphasique, par la persistance de soubassements normatifs à certains marquages diaphasiques ainsi que par une permanence de l'utilisation conjointe de marques et de procédés moins spécifiques pour fournir les informations souhaitées.

1.1 JUSQU'AU MILIEU DES ANNÉES 1980 : DÉFICIT D'EXPLICITATION ET DE COHÉRENCE

Les articles de D. et P. Corbin (1980) et P. Corbin (1989) portaient une appréciation assez critique sur la description de la variété des usages lexicaux par les dictionnaires généraux quelque peu substantiels de l'après-guerre, qui se situait nettement en deçà des typologies de Hausmann.

Une explicitation, plus ou moins aboutie, des principes mis en œuvre ne s'observait que dans le paratexte d'un nombre réduit de répertoires (Corbin 1989, § 2.1) :

- Le mode d'emploi du *Logos* Bordas (1976) répartissait des « marques d'usage » entre « domaine d'emploi » et « niveau stylistique », caractérisait une dizaine de marques référées à celui-ci et spécifiait la portée de chaque emploi d'une marque dans la microstructure des articles (*ibid.* : xij-xiiij).
- Le *Trésor de la langue française* distinguait, parmi les différents « adjuvants démarcatifs de sens » qu'évoquait sa « Préface » (Imbs 1971 : xxxij-xxxiv), le « domaine » et l'« adjuvant stylistique », subdivisé en « niveaux de langue », pour « les situations réciproques de l'auteur et du destinataire de la

3 V. Corbin 1989, § 1 et § 2.1.3 ; 2006 : 44.

4 V., par exemple, le « Glossaire terminologique » de Rézeau 2001 : 16.

5 Le plan du manuel de Gadet (2003), par exemple, se laisse inclure sans difficulté dans le moule hausmannien (v. Corbin 2006 : 44).

communication », et « registres de langue », dédiés à « l'intensité expressive ou affective » (xxxij), sans illustrer concrètement ce principe de classement. Il précisait en outre que des « Remarques » avaient vocation à nuancer, dans les articles, les « indications stéréotypées » fournies par le marquage des mots « “bas” ou “sales” » (xxxiv), mais sa table d'« Abréviations et signes conventionnels » (t. I : cxxvij-cxxxj) n'explicitait pas la valeur de chaque marque.

- Enfin, tous les dictionnaires Robert postérieurs à la première édition du *Petit Robert* (1967) déclinaient, à leur manière et selon leur finalité propre, le classement des « marques d'usage » en fonction des variables « temps », « espace », « société », « fréquence » et « style » opéré dans la « Présentation du dictionnaire » (Rey 1967 : xx-xxj) du répertoire souche, qui, en outre, définissait quelques marques et formulait des réserves sur la validité de « cette notation généralisée des valeurs d'emploi » auxquelles ferait encore écho la « Préface de la deuxième édition » du *Grand Robert* :

Ces marques [...] esquissent – très grossièrement – une configuration des usages de la langue (Rey 1985 : xl)

De plus, les dictionnaires Robert étaient les seuls à proposer dans leurs tables d'abréviations et de symboles une définition pour chaque marque utilisée, tous les autres ouvrages renvoyant tacitement les utilisateurs à leurs propres représentations préconstruites de ce à quoi elles étaient susceptibles de référer.

Quant à la pratique descriptive proprement dite, elle reflétait cet empirisme théorique tant dans le choix des caractérisations des usages que dans leur distribution sur le lexique, et même les dictionnaires les plus explicites concernant leurs principes présentaient à cet égard de grands déficits de cohérence qui empêchaient leur ensemble de marques de constituer un système opératoire, comme D. & P. Corbin (1980) l'avaient montré à propos du *Micro Robert* de 1971 et du *Petit Robert* de 1967.

1.2 PROGRÈS DANS L'EXPLICITATION DES PRINCIPES

Depuis le milieu des années 1980, l'explicitation des principes de description de la variété des emplois s'est consolidée dans les dictionnaires « de langue » Robert, où elle avait déjà une certaine consistance, et elle a tendu à se développer dans les dictionnaires « encyclopédiques » Larousse et Hachette, où elle était quasi inexistante, par la caractérisation de différentes marques dans les tables d'abréviations et de symboles, sans pour autant cependant que leurs préfaces respectives organisent celles-ci en systèmes conceptuels articulés.

- Les dictionnaires Robert généraux ont conservé leurs axes historiques de description des usages, ce dont témoigne le paragraphe « Variétés du français » de la « Préface du *Nouveau Petit Robert* » (Rey-Debove & Rey 1993 : xij-xiv), qui, en tant que cadres des « valeurs sociales d'emploi des mots et des sens », mentionne :

– « le temps »,

– « l'espace », objet d'une insistance particulière au titre des « régionalismes » internes et externes, ces derniers pouvant donner lieu à une codification spécifique s'ils sont des « termes institutionnels » (xij),

- les « *niveaux de langue* », qui, de « l'emploi réservé à la langue écrite et à des discours "soutenus" » aux « contenus qui ne peuvent être exprimés sans danger de choquer » en passant par « les usages qui constituent de véritables signaux d'appartenance sociale » (xiv), recoupent les variables diamédiale, diaphasique, diaévaluative et diastratique de la grille de Hausmann (1989),
- et les « *domaines du savoir* », à cheval sur le diatechnique et le diatextuel :

Certains termes ne sont en usage que dans un domaine particulier de la communication ou dans un type de discours. (*ibid.*)

La cohérence globale des principes descriptifs communs aux différents ouvrages du catalogue se vérifie dans la parenté textuelle des tables d'abréviations et de symboles des refontes respectives du *Grand Robert* (1985), du *Micro-Robert* (1988), du *Petit Robert* (1993) et du *Robert méthodique* (2004, sous le nom de *Robert brio*) comme des nouveaux dictionnaires : *Robert pour tous* (1994), *Robert de poche* (1995, révisé en 2006), *Robert quotidien* (1996), *Robert collègue* (1997) et *Dictionnaire culturel en langue française* (2005). S'agissant de ce dernier ouvrage, qui, sur une matrice adaptée du *Grand Robert*, greffe environ 1 300 articles « culturels » indépendants librement rédigés qui constituent sa spécificité, il est remarquable que les considérations habituelles sur les principes de description de la variation des usages aient été supprimées de la préface (Rey 2005), intégralement axée sur l'originalité du projet.

- Chez Larousse, les préfaces des dictionnaires « encyclopédiques » restent peu portées sur la théorisation : distinction inexplicitée et éphémère entre « niveaux » et « registres » dans le *Petit Larousse* millésimé 1989 (« Aux lecteurs », 6), mais identification entre « niveau de langue » et « registre de communication » dans le *Grand Larousse* de 2005 (« Consulter, utiliser le *Grand Larousse illustré* », viij), et partition entre « niveaux de langue » et « domaines terminologiques » dans le *Dictionnaire général* de 1993 (« Présentation du dictionnaire », vij). Cependant, les tables d'abréviations et signes conventionnels du *Petit Larousse* (éditions de 1989, 1992 et 1998), du *Dictionnaire général* et du *Grand Larousse* explicitent tout ou partie d'un sous-ensemble de 13 des 39 items répertoriés qui peuvent être utilisés comme indicateurs d'usage (distribués sur les 11 types de Hausmann 1989), en affichant des variations textuelles de peu d'ampleur qui attestent à la fois un fond commun et un manque de stabilité. Si on laisse de côté que seuls le *Petit Larousse* 1998 et le *Grand Larousse* mentionnent des items lexicaux auxquels s'appliquent les marques expliquées et que tous les ouvrages sauf le *Petit Larousse* 1989 renvoient vers d'autres marques depuis *vieilli* et *vx*, mais que ces renvois ne sont pas identiques entre le *Dictionnaire général* et les trois autres répertoires, le dispositif se laisse décrire de la façon suivante :
 - des quatre marques explicitées dans les cinq dictionnaires, *didact.* a la même caractérisation dans tous, *vx* voit sa glose plus développée dans le *Petit Larousse* 1989, celle de *vieilli* est légèrement réduite dans le *Petit Larousse* 1998 et le *Grand Larousse*, et *litt.* reçoit deux définitions singulières dans les *Petit Larousse* 1989 et 1998 ;
 - des deux marques non explicitées dans le seul *Dictionnaire général*, *anc.* bénéficie d'une caractérisation affinée dans le *Petit Larousse* 1998 et le

Grand Larousse, cependant que *vulg.* donne lieu à trois variations textuelles, ces deux mêmes dictionnaires étant seuls à l'unisson ;

– *injur.*, non explicitée dans le *Petit Larousse* 1992 et le *Dictionnaire général*, reçoit deux définitions différentes dans le *Petit Larousse* 1989 d'une part et dans le *Petit Larousse* 1998 et le *Grand Larousse* d'autre part ;

– *péjor.*, *sout.* et *très fam.*, qui ne sont explicitées que dans ces deux derniers ouvrages, le sont de la même façon ;

– enfin, *arg.*, *fam.* et *pop.* ne sont caractérisées que par le seul *Petit Larousse* 1989.

Ces distorsions relativisent donc l'avancée que constitue l'explicitation de certaines marques, tout comme le fait que beaucoup demeurent non explicitées dans l'ensemble des ouvrages : pourquoi définir *anc.* et pas *mod.* (utilisées contrastivement, par exemple, *s.v.* ALCADÉ dans tous les répertoires sauf le *Dictionnaire général*), ou encore *litt.* (ex. : ZÉPHYR dans les cinq dictionnaires) et pas *poét.* (ex. : AQUILON, *ibid.*) ? On peut aussi s'interroger sur le fait que, des deux ouvrages qui ont été dérivés en 2003 du *Dictionnaire général*, le *Larousse du collège*, dictionnaire « encyclopédique », hérite des explicitations de marquage du répertoire source, mais pas le *Larousse pratique*, dictionnaire « de langue ».

- Parallèlement à cette évolution des dictionnaires « encyclopédiques » Larousse, le millésime 1994 du *Dictionnaire Hachette encyclopédique* associa similairement une définition aux 19 éléments réunis dans sa table des « Marques d'usage » (xiv), qui couvraient tous les types de Hausmann (1989) sauf le dia-intégratif, mais n'étaient référés explicitement qu'aux trois variables « lieu », « temps » et « faits de société ». Lors de la refonte millésimée 2002, cet inventaire (xij) a été réduit à 16 unités (par suppression des marques diatopiques *régional* et *dialectal* et de la marque diafréquentielle *rare*), dont une moitié présentent une évolution définitionnelle.
- Par contre, les concurrents mineurs du *Petit Larousse* (l'éphémère *Maxidico* en 1996, le *Dictionnaire encyclopédique Auzou* en 2002) ne se sont pas signalés de façon remarquable sur le terrain de la description de la variété des usages.

1.3 GLISSEMENTS DU DIASTRATIQUE AU DIAPHASIQUE

Dans le *Petit Robert* de 1967 et le *Micro Robert* de 1971, la variable « société », qui précédait les « niveaux de langue » du *Nouveau Petit Robert* de 1993 (v. *supra* § 1.2), était au cœur des dysfonctionnements observés par D. & P. Corbin (1980, v. *supra* § 1.1), en particulier au titre des flottements auxquels donnaient lieu les marques *familier* et *populaire*. À partir des années 1980, les dictionnaires Robert ont connu en la matière, sous l'influence d'Alain Rey, une évolution positive dont on trouve les prémices dans le titre conservatoire d'un répertoire spécialisé dont il fut coauteur : le *Dictionnaire du français non conventionnel* (Cellard & Rey 1980). Une théorisation suivit, dans un article (Rey 1983) qui stigmatisait le fait que « la marque "populaire" signifiant "unité employée par des locuteurs appartenant au peuple" » fût « trop souvent utilisée comme un intensif de "familier", qui devrait vouloir dire : "unité convenant à un type de communication non officielle, quotidienne, sans contrainte, quelle que soit l'appartenance sociale des communicants" ». L'auteur appelait à « distin-

guer les marques d'appartenance sociale », dont l'« utilité n'est pas évidente », « des marques de situation communicative, comme “familier”, qui seront à opposer à “soutenu” et non pas à “littéraire” qui, sur le même axe, spécifie que la communication est écrite et aboutit à un texte fonctionnant en tant que “littérature” dans la société » (*loc. cit.* : 564). Les effets dictionnaires corrélés de cette réflexion furent, deux ans plus tard, la régression de la marque « populaire » et l'expansion de la marque « familier » dans la refonte du *Grand Robert*, explicitées avec netteté dans la « Préface de la deuxième édition » :

L'abréviation *pop.* pour « populaire » a été beaucoup moins utilisée que dans la première édition du dictionnaire, et même que dans le *Petit Robert*. C'est qu'on a renoncé à qualifier de « populaires » des mots et des emplois que toute la communauté employait dans certaines circonstances de la communication. [...] Ainsi, « populaire » sera dans ce dictionnaire réservé aux emplois réprouvés par les personnes pourvues d'un « capital scolaire » (*de suite pour tout de suite* [...] ; *votre dame pour votre femme*, etc.). Au contraire, l'abréviation *fam.* (familier) est ici très employée. Elle correspond aux emplois (formes et sens) normaux dans une communication sociale aisée, plutôt parlée qu'écrite, et dénuée de la contrainte propre aux échanges officiels, hiérarchiques, etc. Elle peut se moduler en « très familier », en « familier et vulgaire », quand le contenu (érotique, scatologique, etc.) est en cause [...]. (Rey 1985 : xl)

Cette option, qui lève une ambiguïté mais crée une difficulté nouvelle en codant comme des degrés (*familier, très familier*) des appréciations qui ressortissent respectivement, selon les caractérisations mêmes qui en sont données, au diaphasique et au diaévaluatif (v. *infra* § 2.1), se propagea progressivement dans le catalogue Robert et trouva en 1993 chez Larousse, avec le *Dictionnaire général*, un pendant qui, à défaut de satisfaire pleinement aux exigences de l'adéquation descriptive de la variation des usages, témoigne de l'évolution des sensibilités dans l'appréciation des critères pertinents pour en rendre compte :

Nous avons exclu la marque POP. (populaire), encore présente dans de nombreux dictionnaires, la jugeant inadaptée au français d'aujourd'hui. Cependant, nous avons introduit une gradation dans le registre familier FAM. et T.FAM. (très familier) – qui nous paraît plus pertinente pour guider le lecteur dans son usage de la langue. Inversement, à côté de la langue dite littéraire, nous avons introduit le niveau SOUT. (langue soutenue), qui rend compte d'une langue maîtrisée et choisie, à l'oral ou à l'écrit. (« Présentation du dictionnaire », p. vii)⁶

1.4 FLOTTEMENTS ENTRE DIAPHASIQUE ET DIANORMATIF

Une autre hypothèque pèse sur la marque *familier*, qui tient au fait que l'ensemble de ses utilisations, même dans les dictionnaires récents précédemment évoqués qui ont fait évoluer leur conception du marquage des usages, ne se laisse pas ramener à la variable diaphasique assumée par Rey (1983 et 1985), mais que beaucoup d'entre elles, au lieu de répondre à la vocation socio-énonciative affichée, servent à fustiger des propriétés formelles, combinatoires ou sémantiques d'unités lexicales qui résultent de mécanismes fondamentaux de l'évolution du lexique et de son adaptation aux besoins des locuteurs. La place manquant pour donner la mesure exacte de cette dimension normative du

6 Pour une version plus développée de ce qui est exposé dans ce paragraphe, v. Corbin 2006 : 48-49.

marquage présenté comme diaphasique (dont la première hypothèse fut formulée par Corbin 1989, § 3.2.2), nous n'évoquerons que l'exemple parlant du mot ACCIDENTÉ, adjectif et nom, marqué « Fam. » dans la deuxième édition du *Grand Robert* (et encore dans le *Nouveau Petit Robert*) quand il s'applique à des personnes et des véhicules, probablement contre l'intuition de nombre de locuteurs et en dépit de son emploi commun dans des environnements énonciatifs sérieux, voire officiels⁷. La cause de ce marquage, qui n'a rien à voir avec les conditions effectives d'utilisation du mot, est à trouver dans le déplacement de l'usage de celui-ci : alors que son application première à des paysages (*relief accidenté, région accidentée*) n'est pas marquée, son emploi plus récent mais bien installé à propos d'êtres humains (*personne accidentée, accidenté de la route, du travail*) et de véhicules (*voiture accidentée*) a en effet fait l'objet de critiques puristes⁸, dont le *Grand Robert* se fait l'écho par la mention légitime « (emploi critiqué) », que la marque « Fam. » qui la précède, dévoyée de sa vocation diaphasique affirmée, ne fait que redoubler⁹. Cet exemple donne une idée du chemin qui reste à parcourir pour désidéologiser le marquage diaphasique et mettre en la matière les pratiques en accord avec les principes.

1.5 MARQUAGE AVEC ET SANS MARQUES

Aussi bien Hausmann (1977 : 115-116) que Corbin (1989, § 3.1.2) avaient observé que les informations sur la spécificité des usages pouvaient non seulement être exprimées de manière codifiée par des « marques » autonomes mais aussi prendre une forme textuelle moins réglée, en s'intégrant par exemple aux définitions. Si le marquage codé répond à des visées et à des contingences de la lexicographie moderne (cohésion textuelle, structuration informatique des articles, économie typographique), la manière rédactionnelle héritière de la lexicographie des origines perdure cependant dans les dictionnaires généraux actuels, comme en témoigne, par exemple, la conjonction de la marque générale « PHYS. » et du segment définitionnel plus spécifique « en physique des particules » dans le *Nouveau Petit Robert s.v.* 1. EFFICACE 1. : « PHYS. *Intensité, tension efficace* : valeur moyenne de l'intensité, de la tension d'un courant alternatif, équivalente à celle d'un courant continu. *Section efficace* : en physique des particules, grandeur utilisée pour rendre compte d'une expérience de diffusion ». La préface de ce dictionnaire affiche même l'intention d'utiliser à des fins particulières chacun des deux modes textuels pour certains marquages diatopiques :

Pour les mots et usages propres à ces pays et territoires, on précise la zone d'emploi (« Québec », « français d'Afrique », etc.). Les termes institutionnels, pour leur part, sont distingués lorsqu'ils n'ont cours qu'en français de France, ou

7 Voir les sites de diverses institutions ou entreprises françaises : Fédération nationale des accidentés du travail et des handicapés, sécurité sociale, sécurité routière, ministères de la santé, de l'emploi, des transports, compagnies d'assurances, etc.

8 Elles sont détaillées dans Dupré (1972), *s.v.* ACCIDENTÉ.

9 Pour mémoire, cet emploi d'*accidenté*, qui remonte à 1900 selon le *Grand Robert* et que la 9^e édition du *Dictionnaire de l'Académie française* marque « Fam. et abusivt », n'est plus marqué par le *Petit Larousse* depuis le millésime 1981, après avoir été affecté successivement des marques « Abusiv. » lors de son insertion dans le dictionnaire (millésime 1938), puis « Fam. » à partir du millésime 1960.

en français de Suisse ou de Belgique, ou du Canada francophone, par la mention Au Québec, En Suisse, etc. (Rey-Debove & Rey 1993 : xijj)

Mais la mise en pratique semble quelque peu aléatoire : si l'article CÉGEP, qui est affecté d'un marquage de ce dernier type, réfère bien à une entité institutionnelle (« Au Québec, Collège d'enseignement général et professionnel, situé entre le secondaire et l'université »), il n'en va pas de même pour ÉRABLIÈRE, dont le marquage est cependant exprimé de la même façon (« SPÉCIALT Au Québec, Plantation d'érables à sucre, dans la partie limitrophe boisée d'une propriété agricole, exploitée pour l'industrie des produits de l'érable »). Le *Petit Larousse*, qui ne tient pas, en la matière, de discours explicite sur les formes du marquage, recourt en pratique aux deux mêmes types textuels pour exprimer des spécificités diatopiques, selon une répartition qui, elle aussi, suscite des interrogations : « Québec. Peuplement d'érables exploité en acériculture. » (s.v. ÉRABLIÈRE 2. depuis le millésime 2005) contraste avec « Au Québec, exploitation d'une érablière pour produire du sirop d'érable et d'autres produits dérivés. » (s.v. ACÉRICULTURE dans le millésime 2003, substitué à « Québec. », qui, depuis le millésime 2000, remplaçait « Canada. »).

2. DICTIONNAIRES POUR APPRENANTS

Alors que l'offre dictionnaire à l'intention des utilisateurs supposés confirmés présente une certaine stabilité dans la durée, la période récente a connu un renouvellement important des répertoires destinés aux apprenants étrangers et surtout natifs, dans lesquels les indications afférentes à la variété des usages lexicaux trouvent des expressions différentes en fonction des finalités propres à chaque type d'ouvrage.

2.1 FRANÇAIS LANGUE ÉTRANGÈRE

Vingt ans après le *Dictionnaire du français langue étrangère* (niveaux 1 et 2) de Larousse, le *Dictionnaire du français* (Le Robert et CLE International) est venu combler en 1999, pour les allophones apprenant le français, le vide laissé par la disparition de son prédécesseur, mais à une échelle de nomenclature supérieure (22 000 entrées contre 10 000 mots traités ou évoqués dans 5 000 articles) et avec un concept dictionnaire différent de celui-ci, qui fait une place importante à la variation des usages dans son paratexte comme dans les articles.

L'« Avant-propos » spécifie la teneur de la nomenclature du dictionnaire :

Ce sont les mots courants de la conversation et de la presse. Les termes techniques et scientifiques des professionnels ont été écartés. (Rey-Debove 1999 : vij)

Une hiérarchie est établie entre les mots « qu'il faut connaître en priorité » et les autres, fondée sur « la notion d'importance » (ix), qui est substituée aux données brutes de la fréquence :

Est *important* un mot dont on ne peut se passer dans la vie quotidienne, ou pour comprendre, ou pour s'exprimer.

et qui transcende les distinctions de registres :

Il arrive souvent qu'un synonyme familier soit beaucoup plus employé que le mot neutre (*P.-V.* est plus courant que *contravention*, *flic* plus courant que *agent* ou *policier*), et on en a tenu compte.

Les « niveaux de langue », justement, correspondent fondamentalement à trois « registres » (vij) diaphasiques : « normal, style familier, style recherché », le *style familier* s'observant dans « des situations de spontanéité amicale » et le *style recherché* dans le « vocabulaire des gens cultivés ou qui soignent leur style » (xj). Mais la « Liste des termes, signes conventionnels et abréviations » (xij) mâtine celui-là de diamédial et de diafréquentiel (« emploi essentiellement oral d'un mot courant »), celui-ci de diafréquentiel seulement (« emploi d'un mot un peu plus rare et plus élégant que ses synonymes »), et la gradation de *familier* à *très familier* fait, comme dans le *Grand Robert* (v. *supra* § 1.3), glisser du diaphasique au diaévaluatif (« met en garde contre des emplois trop négligés ou vulgaires » (Rey-Debove 1999 : xj), « dont la signification peut être choquante ou agressive » (« Liste ... », xij). Et cette dernière variable peut aussi être prise en charge par la rubrique de remarques de certains articles :

Lorsqu'un mot est dangereux (discourtois, péjoratif, « politiquement incorrect »), on le précise en remarque. (Rey-Debove 1999 : xj)

Dans sa mise en pratique, ce dispositif, qui, didactiquement, signale les mots ou emplois « importants » au moyen d'« une *flèche bleue* placée en marge » (*ibid.* : ix) et formule sans abréviation les marques codifiées, couvre en fait un éventail plus large des variables de Hausmann (1989) que ce que laissent attendre les paratextes. La rubrique de remarques peut notamment s'ouvrir

- au diachronique : « On peut écrire aussi *phantasme*, mais cette orthographe est vieillie. » (s.v. FANTASME) ;
- au dia-intégratif : « *Farniente* est un mot italien que l'on prononce aussi [farnjente] en français. » (s.v. FARNIENTE) ;
- au diastratique : « Ce mot est surtout employé par des hommes. » (s.v. GONZESSE) ;
- au diatechnique : « En psychiatrie, on ne parle plus de *folie* mais de *maladie mentale* ou de *troubles mentaux*. » (s.v. FOLIE) ;
- au dianormatif : « Cette prononciation est jugée familière par les puristes. » (s.v. 2. FAIT) ;
- ou tout à la fois au diatopique et au dianormatif (« Ce mot est employé au Québec, pour éviter l'anglicisme *shopping* que l'on utilise en France. » s.v. MAGASINAGE) ou au diatextuel et au diastratique (« *Épouse* est un terme administratif ou populaire. » s.v. FEMME).

Et les définitions sont susceptibles, comme dans les dictionnaires généraux, de relayer les marques en certaines occasions, par exemple au titre d'indications diatopiques (« Dans le midi de la France, Un peu fou, cinglé. » s.v. FADA). Le marquage, qui est abondant, s'étend aussi opportunément aux renvois, comme il est habituel dans les dictionnaires Robert. Cependant le nombre élevé de ceux-ci vers des items usuels mais ressortissant à des variétés d'usages connotées négativement peut surprendre dans la perspective d'aide à l'expression qui sous-tend ce dispositif analogique, puisqu'il oriente les allophones vers des items dont l'emploi peut être « dangereux » : « → *agaçant* ; style familier *chiant*, *embêtant*, *emmerdant*. » (s.v. FATIGANT 2.). Par ailleurs, trois décennies après le *Micro-Robert* de 1971, le *Dictionnaire du français* manifeste la même incapacité que celle qu'y ont relevée D. & P. Corbin (1980, § « Les renvois dans le *Micro-Robert* », v. *supra* § 1.1) à affecter un marquage constant aux

différentes occurrences d'un même item dans le dictionnaire, ce qui ne peut qu'être un facteur de trouble pour les utilisateurs :

- parmi les items qui viennent d'être mentionnés, *chiant* est marqué « STYLE TRÈS FAMILIER » à son entrée propre ou comme renvoi s.v. EMMERDANT 2. et *embêtant* est non marqué en tant que renvoi s.v. EMMERDANT 1. ;
- et il est aisé de trouver d'autres exemples : *flapi*, non marqué à son entrée et « style familier » s.v. FLAGADA ; *bluff*, non marqué à son entrée et comme renvoi s.v. FRIME mais « style familier » comme renvoi s.v. ESBROUFE ; *dégonflé*, « style familier » à son entrée (2.) mais non marqué comme renvoi s.v. TROUILLARD 2. ; *foutre*, « style très familier » à son entrée mais seulement « style familier » comme renvoi s.v. FABRIQUER 3. ; etc.

Paru trois ans après le *Dictionnaire du français*, le *Dictionnaire du français usuel* (Picoche & Rolland 2002) publié chez De Boeck - Duculot, conçu sur de tout autres bases, n'apporte pas d'innovations remarquables en matière de caractérisation des usages. Des listes de fréquence existantes étayent la sélection de la nomenclature, voulue la plus usuelle (v. « Présentation et mode d'emploi », p. 7-11). Pour le reste, la prise en charge des faits de variation se limite à une énumération des marques retenues sans autre commentaire, dont l'originalité réside dans l'utilisation de la marque *sav.* (« savant »), sans que sa valeur soit pour autant explicitée :

Nous jouons sur plusieurs niveaux de langage et nous les signalons, généralement par des abréviations dont la table figure ci-dessous. Ainsi, *litt.* (littéraire), *sav.* (savant), *fam.* (familier), *vulg.* (vulgaire), *argot* et *techn.* quand il nous arrive de signaler un mot d'argot ou un mot technique d'usage courant. (*ibid.* : 8)

2.2 CYCLE 3 ET FRONTIÈRE ÉCOLE-COLLÈGE

Avec une petite dizaine de titres, l'offre dictionnaire faite aux élèves du cycle 3 de l'école élémentaire ou entrant au collège est plus diversifiée qu'à destination des apprenants allophones « adolescents ou adultes » (Rey-Debove 1999 : vij), réduite au seul *Dictionnaire du français*, et si les nomenclatures sont commensurables, elles présentent des différences qualitatives importantes et significatives. Loin de l'image non édulcorée des usages lexicaux à laquelle vise ce dernier (v. *supra* § 2.1), les répertoires pour jeunes natifs se concentrent sur un lexique fondamental de bon aloi utile scolairement, comme en témoigne la « Préface » du *Petit Robert des enfants*, qui eut le même maître d'œuvre que le *Dictionnaire du français* :

On a fait l'économie des mots très familiers au profit d'un vocabulaire utile et moins connu. Et, quant aux mots plus rares, les enfants les acquerront progressivement avec l'âge, l'essentiel étant d'avoir commencé par une bonne maîtrise du lexique de base et de son fonctionnement. (Rey-Debove 1988 : x)

La place contrôlée faite à la variation diaphasique peut, comme c'est le cas dans le *Maxi débutants* Larousse de 1997, être présentée comme un point d'appui pour l'apprentissage du lexique châtié :

Le rôle du dictionnaire est surtout de t'apprendre le langage soigné, et si beaucoup d'expressions et de mots familiers s'y trouvent, c'est pour que tu remarques bien l'abréviation *fam.* et que tu apprennes comment on peut dire la même chose en langage soigné. (« À la découverte du "Maxi débutants" », ix, en gras dans le texte source)

Dans le *Larousse junior* de 2008 (plus explicite que l'édition originale de 2003), le propos, sous l'intitulé « Les registres de langue », est voisin mais avec une dominante diamédiale :

comme on n'utilise pas toujours les mêmes mots selon que l'on écrit ou que l'on parle, ce dictionnaire précise les mots qui sont *familiers* (ceux que l'on utilise avec ses camarades) ainsi que ceux qui appartiennent à la langue *littéraire* (ceux que l'on trouve surtout à l'écrit). (« À la découverte du *Larousse junior* », xj)

Par ailleurs, la nomenclature des dictionnaires pour jeunes apprenants peut connaître des infléchissements diatechniques suscités par le vocabulaire des diverses matières scolaires :

Un certain nombre de mots spécifiques aux autres disciplines enseignées, et utilisés notamment dans les manuels scolaires (sciences et techniques, histoire, géographie...), sont également traités afin d'enrichir encore le vocabulaire. (« Avant-propos » du *Larousse junior* de 2003, v ; texte voisin dans l'édition de 2008, vij)

et s'ouvrir à la littérature de jeunesse :

Nous n'avons pas négligé le vocabulaire que l'on trouve dans les textes littéraires courants du C.M.2 et de la classe de 6^e. Par exemple : *famélique*, *fugace*, *draconien*, *besace*, *capiteux*, *chapon*, *chimère*, *escarcelle*, *arborer*, *élimé*, *cynique*, *infortuné*, etc. (« Préface » du *Hachette junior* de 1986, p. 3)

Ces options quant au choix des nomenclatures convergent avec l'adaptation du texte des dictionnaires à l'âge des élèves pour simplifier l'appareil de description de la variété des usages. Les paratextes sont peu diserts, le discours spécifique des préfaces étant presque intégralement contenu dans les extraits qui viennent d'être cités, et les modes d'emploi, quand ils existent, se limitant, à l'exception de ceux des deux dictionnaires Larousse évoqués ci-dessus, à pointer l'indication qu'un mot est « familier » dans un article modèle (*Hachette junior(s)*, *Robert junior* et *Super major* Larousse depuis leur première édition). Quant aux tables de signes conventionnels, qui, elles non plus, ne sont pas toujours présentes, elles ne renseignent sur des marques susceptibles d'être utilisées que pour autant que les dictionnaires optent pour leur abréviation, ce qui n'est le cas que pour un sous-ensemble des répertoires et, dans les éditions les plus récentes, ne vaut plus que pour le *Maxi débutants* et *Mon grand dictionnaire de français* : sur le dernier quart de siècle, tous ouvrages et toutes versions confondus, on ne rencontre alors que *fam.*, *très fam.*, *littér.* et *péjor.*

Par le jeu combiné, déjà observé dans le *Dictionnaire du français*, des marques proprement dites, des définitions et des remarques, les dictionnaires pour le cycle 3 et le début du collège couvrent quand même, dans des proportions différentes et avec des variations d'un ouvrage à l'autre qui mériteraient qu'on s'y appesantisse, la quasi-totalité des variables de la grille de Hausmann (1989) :

- Le diaphasique domine de façon générale avec le très usité *familier* mais s'exprime aussi, assez communément, dans les remarques du *Super major* depuis sa première édition (1994), sous la forme « Ce mot / Cette expression appartient à la langue soutenue. » (par ex. *s.v.* FATUITÉ ou GOUVERNER). Avec l'usage non explicité de *littéraire* (ex. : FANGE dans le *Hachette junior(s)* depuis 1986 et le *Scolaire Hachette* (2002) qui le prolonge vers le collège, le

Larousse junior de 2003, le *Robert junior* de 2005, le *Auzou junior* (2006)), le diaphasique se différencie mal du diatextuel, qui par contre s'actualise sans ambiguïté dans le *Super major* dans la formule de remarques « Ce mot appartient à la langue poétique. » (ex. : FIRMAMENT). Pour des données de même nature, le *Maxi débutants* de 1997 préfère en revanche un axage diamédial : « Ce mot s'emploie surtout dans la langue écrite. » (en remarque s.v. FANGE). À cheval sur le diamédial et le diaphasique, on peut aussi évoquer les indications affectant des mots propres aux échanges entre adultes et enfants en bas âge : « Ce mot est utilisé par les très jeunes enfants. », remarque de l'article JOUJOU du *Robert junior* de 1993 corrigée en « On utilise ce mot quand on s'adresse à de très jeunes enfants. » dans l'édition de 2005.

- Le diachronique n'est pas absent, exprimé directement par des marques dans le *Larousse junior* dès 2003 (« Mot ancien » s.v. FAUBOURG) ou, plus souvent, dans des définitions. Il peut alors être pris en charge par des indicateurs temporels : « autrefois » pour ESCARCELLE dans le *Hachette junior(s)* de 1986 à 1993 inclus, le *Maxi débutants* depuis 1986, le *Robert junior* et le *Super major* dès l'origine (1993), *Mon grand dictionnaire illustré* (2000) et son successeur le *Grand dictionnaire de français* (2004) ; « se disait autrefois » pour GUEUX dans le *Maxi débutants* depuis 1986 ; « terme ancien » pour GUEUX dans le *Super major* et pour GAZETTE dans le *Auzou junior* ; « vieilli » pour GAZETTE dans le *Hachette junior* depuis 1998 et le *Scolaire Hachette* ; « un peu vieilli » pour GREDIN dans le *Super major*. Mais il arrive aussi qu'il soit seulement suggéré par le temps verbal : par exemple, pour ESCARCELLE, l'imparfait dans le *Scolaire Hachette* (« Grande bourse que l'on suspendait à la ceinture. ») et le *Auzou junior*. Par ailleurs, les remarques peuvent également être mises à contribution : « Ce mot ne s'emploie plus beaucoup, on dit aujourd'hui "employé de maison". » (s.v. 2. DOMESTIQUE dans le *Super major*).
- L'information diatopique privilégie les définitions et les remarques :
 - Les premières sont particulièrement utilisées dans le *Super major*, dès 1994 (« En Suisse » pour GYMNASE 2, « en France » pour le garde des Sceaux s.v. 1. GARDE, « En Alsace, dans le Nord et en Belgique » pour AUBETTE, « en Camargue » pour GARDIAN) ou seulement depuis 1997 (« Au Canada » pour GLISSOIRE, « au Québec » pour gomme à mâcher s.v. GOMME).
 - Les secondes sont mises à contribution dans ce même dictionnaire à partir de 1994 (« Ce mot s'emploie surtout en Bretagne et en Normandie. » s.v. GOÉMON) et dans le *Robert junior* de 2005, qui y redistribue des indications prédéfinitionnelles de ses éditions antérieures (« Ce mot est employé en Belgique. » s.v. AUBETTE, « Ce mot est utilisé en Suisse. » s.v. FOURRE, « Ce mot est utilisé dans le Nord de la France et en Belgique. » s.v. WASSINGUE).

Les francophonismes paraissent l'emporter sur les régionalismes dans les deux ouvrages, qui ne font cependant pas la même part aux canadianismes, privilégiés dans le premier et, semble-t-il, absents du second.
- Pour les mots et les sens propres à des emplois spécialisés, divers dictionnaires contournent par les définitions ou les exemples la spécification de cette spécialisation, que quelques autres explicitent par l'insertion d'un indicateur diatechnique dans la définition : « En physique » (s.v. FORCE 1.), « En

mathématiques » (s.v. FRACTION 2.), « Au théâtre » (s.v. 2. GÉNÉRALE), « En botanique » (s.v. 1. GREFFE 1.) dans le *Super major* depuis 1994, « Au cinéma » (s.v. FLASH-BACK) dans ce dictionnaire et dans le *Auzou junior*, « En grammaire » (s.v. GROUPE 5. ou 4.) dans ces deux ouvrages, le *Larousse junior* dès 2003 et le *Robert junior* de 2005.

- La variable dia-intégrative établit également un partage tranché entre dictionnaires, les uns évitant de signaler les emprunts comme tels (*Nouveau Larousse des débutants* (1977) puis *Maxi débutants*, *Auzou junior*) alors que d'autres, indépendamment d'éventuelles précisions proprement étymologiques, font état de l'origine de certains d'entre eux pour l'articuler à des données à caractère fonctionnel, au premier rang desquelles figurent des indications de prononciation¹⁰ : ainsi en va-t-il dans le *Super major* depuis 1994 et le *Larousse junior* depuis 2003, qui dupliquent dans divers articles le type de remarque que l'on trouve par exemple s.v. GROOM (« C'est un mot anglais, on prononce [grum]. »), le *Hachette junior* depuis 1998 et le *Scolaire Hachette*, qui intègrent dans leur rubrique de prononciation des données comparables (« Groom est un mot anglais : on prononce [grum]. »), ainsi que le *Fleurus junior* (2001), qui conjoint avec systématisme les deux types d'informations (« Mot anglais qui se prononce [gol]. » s.v. GOAL). Il peut également advenir que l'information dia-intégrative soit associée à une observation orthographique, comme s.v. STICK dans le *Larousse junior* dès 2003 (« C'est un mot anglais. – Il se termine par ck. ») ou TOMAHAWK dans le *Robert junior* depuis 2005 (« Ce mot indien s'écrit aussi tomawak. »).
- C'est aussi dans des remarques que s'expriment volontiers, dans certains répertoires, des notations diafréquentielles occasionnelles (« S'abrège souvent en *frigo*. » s.v. FRIGIDAIRE dans le *Super major* dès 1994, « Le féminin *grognonne* est rare. » dans le même dictionnaire et le *Maxi débutants* de 1997 s.v. GROGNON, « il s'emploie très peu » dans le *Robert junior* s.v. GROGNON depuis 2005) et dianormatives (« Il est recommandé d'employer l'expression *retour en arrière*. » s.v. FLASH-BACK dans le *Super major* depuis 1997, « On dit plutôt *les malades mentaux*. » en remarque postdéfinitionnelle s.v. FOU 1. dans le *Robert junior* dès 1993). Mais il peut arriver qu'une information sur la fréquence soit prise en charge par la définition (« *fange* n.f., *fangeux* adj. sont des équivalents rares de *boue*, *boureux*. » dans le *Maxi débutants* de 1986).
- Le marquage diaévaluatif, quant à lui, quand il n'est pas codé par une marque (« Très fam. » dans le *Maxi débutants* depuis 1986 pour GUEULER, « (Très familier) » pour ce même mot (sens 1.) et GUEULE 2. dans le *Super major* de 2004), trouve volontiers son expression au sein des définitions (« Mot affectueux » s.v. FILLETTE 2. dans ce même dictionnaire dès 1994, « Dans le langage très familier » s.v. GUEULE 2. dans le *Auzou junior*). Mais il peut aussi faire l'objet d'une remarque (« *Gueule* est un mot grossier quand on parle d'une personne. » dans le *Larousse junior* de 2003¹¹). L'article

10 V. Gasiglia 2008a ; 2008b, § 3 ; 2008c, § 2.

11 Remarque affinée en « *Gueule* est un mot grossier quand on parle de la bouche ou du visage d'une personne. » dans l'édition de 2008.

NÈGRE, qui donne lieu à un marquage généralisé comme « péjoratif » et/ou « raciste », voit celui-ci formulé des trois manières possibles selon les répertoires : par une marque (*Hachette junior(s)* jusqu'à l'édition de 1993 incluse, *Larousse junior*), dans une définition (*Nouveau Larousse des débutants* et *Maxi débutants*, *Hachette junior* depuis 1998, *Scolaire Hachette*, *Fleurus junior*, *Grand dictionnaire de français*, *Auzou junior*) ou dans une remarque (*Robert junior* de 1993 et 1999, *Super major*).

Dans les définitions et les remarques, il n'est pas rare que la recherche d'adéquation dans la description de la variété des usages fasse conjointre des éléments de marquage relevant de plus d'une variable. Le dianormatif est ainsi associé

- au diachronique et au diaévaluatif dans la remarque de l'article NÈGRE du *Robert junior* de 2005 : « Ce mot est vieux ou péjoratif. On dit maintenant *un Noir, une Noire* ou *une personne de couleur*. » ;
- au diafréquentiel dans celle de l'article GÂCHETTE du *Super major* : « On dit souvent, à tort, *appuyer sur la gâchette* au lieu de *appuyer sur la détente*. » ;
- au dia-intégratif pour différents anglicismes dans ce même dictionnaire et dans le *Maxi débutants* de 1997 (« Ce mot est la forme française de l'anglais *gas-oil*. » s.v. GAZOLE) et, sous une forme plus marquée, dans le *Larousse junior* depuis 2003 (« C'est un mot anglais, il vaut mieux dire *gardien de but*. » s.v. GOAL) ;
- ou encore au diatopique dans les remarques de l'article GROOM du *Super major* depuis 1997 : « 1. C'est un mot anglais on prononce [grum]. 2. Au Québec, on doit dire *chasseur*. »

Et la variable diatopique elle-même peut également, dans ce même dictionnaire, se trouver combinée avec une information diafréquentielle (« Ce mot est rare au Québec. » en remarque de GROGGY) ou diaphasique (« Mot familier qu'on utilise surtout dans le sud de la France » dans la définition de FADA depuis 1994).

2.3 AUTOUR DU CYCLE 2

Le changement d'échelle (entre 6 000 mots et un millier) et de nature des dictionnaires, scolaires ou non, qui encadrent l'âge de l'apprentissage de la lecture et de l'écriture fait qu'on ne saurait y retrouver la relative richesse de l'information sur la variété des usages observable dans les répertoires pour le cycle 3 de l'école primaire. Pour autant, celle-ci, quoique de portée restreinte et limitée à certains ouvrages, n'est pas inexistante, et elle présente même certaines spécificités.

- Le créateur du genre, le *Mini débutants* Larousse de 1985, comportait, dans son composant microstructurel indifférencié, quelques indications éparées, référant un petit nombre de mots à l'univers textuel des contes et des poésies (FÉE, LICORNE, MARÂTRE, OGRE, OR (conjonction), SIRÈNE), y ajoutant parfois une note diamédiale (« Ce mot se trouve surtout dans les contes, les histoires écrites » s.v. DEMEURE), donnant occasionnellement des repères diachroniques par des moyens lexicaux (« Autrefois » s.v. QUENOUILLE, « Maintenant, on dit plutôt *agriculteur*. » s.v. PAYSAN) ou le seul jeu du temps verbal (« C'était un domestique... » s.v. VALET), et exceptionnel-

lement un indice diaphasique (« C'est un autre mot, qu'on dit entre nous, pour *ami*, *camarade*. » s.v. COPAIN, v. Corbin 1989, § 3.1.1) ou diaévaluatif (« C'est un petit mot gentil » s.v. GAMIN).

- Une décennie plus tard, postérieurement à l'instauration du cycle 2 de l'école primaire, différents dictionnaires de format et de conception variables sont venus concurrencer le *Mini débutants*. Ses deux rivaux les plus comparables, apparus au milieu des années 1990, ont retenu des options différentes :
 - Le *Hachette benjamin* (1996) a sélectionné une nomenclature peu sujette au marquage et s'est abstenu d'indications sur les conditions d'emploi de mots qui auraient pu en recevoir (ex. : COPAIN, EMBÊTER, GOSSE, etc.).
 - Le *Robert benjamin* (1997) a suivi, de façon aussi mesurée mais en utilisant tour à tour les définitions et les remarques, le modèle du *Mini débutants* sur les terrains diatextuel (« C'est un mot que l'on trouve dans les contes de fées » s.v. DEMEURE), diamédial (« C'est un mot que l'on ne dit pas beaucoup, on le rencontre surtout dans les livres. » s.v. FLOT), diachronique (« Maintenant on dit plutôt un *agriculteur*. » s.v. PAYSAN, « Aujourd'hui, on dit plutôt *employé de maison*. » s.v. DOMESTIQUE) et diaphasique (« Embêter et s'embêter sont des mots que l'on dit seulement avec ses amis ou sa famille. » s.v. EMBÊTER, « Quand on parle avec des gens sérieux, c'est mieux de dire un *petit garçon* ou une *petite fille*. » s.v. GAMIN) ; en outre, il y a ajouté sporadiquement des observations diafréquentielles (« Or est un mot que l'on n'emploie pas beaucoup. On dit plutôt *cependant* ou *mais*. » s.v. OR) parfois associées à des prescriptions dianormatives (« On dit souvent *appuyer sur la gâchette*, mais on a tort. Il faut dire *appuyer sur la détente*. » s.v. GÂCHETTE).
- Par rapport à son édition première et à ses concurrents, le *Mini débutants* Larousse s'est ensuite renouvelé en deux temps :
 - Le *Dictionnaire Mini débutants* de 1999 a présenté deux évolutions significatives : d'une part l'ajout à sa nomenclature d'un petit ensemble de mots liés de façon privilégiée à un univers de textes pour enfants (API, BARBICHETTE, CHAPERON, CHOIR, ÉLIXIR, ENCHANTEUR (nom), GRIFFON, LIEUE, MATINES, VIZIR...) et illustrés d'une citation de l'un de ceux-ci (« *On lui fit faire un petit chaperon rouge, qui lui allait si bien, que partout on l'appelait le Petit Chaperon rouge*. » le Petit Chaperon rouge », conte de Perrault » s.v. CHAPERON) ; d'autre part, pour le marquage diaphasique, l'introduction concertée de la marque *familier* en remarque dans la formulation récurrente « Attention ! C'est un mot familier. », appliquée le plus souvent à des mots nouvellement accueillis (BALADE, BIDULE, COSTAUD, EMBÊTER, FROUSSE, GOINFRE, GOSSE, LOUPER, MACHIN, MOCHE, PAGAILLE, PATATE, RAFFUT, RÂLER, RIGOLER, ROUSPÉTER...).
 - Sous le titre de *Larousse des débutants*, l'édition de 2005 prolonge cette orientation et se singularise par la substitution aux citations d'une icône placée avant les adresses concernées qui « indique un mot que tu trouves dans un conte, une chanson » (« Comment utiliser le *Larousse des débutants* », vj) et par la suppression de « Attention ! » en introduction des remarques.

- Parmi les autres répertoires, relativement nombreux, destinés au même public, qui s'écartent du modèle partagé par les dictionnaires des trois éditeurs majeurs pour le cycle 2, seuls deux ouvrages se distinguent :

– Le *Petit Fleurus* (1998), dont le texte et l'iconographie s'ancrent dans l'univers des contes, fournit occasionnellement, dans une définition ou en remarque, des indications diachroniques (« Autrefois » s.v. LIEUE), diafréquentielles (« C'est aussi le nom qu'on donne parfois aux poules. » s.v. COCOTTE) ou, indirectement, diatextuelles (« Les *licornes* n'existent que dans les légendes. » s.v. LICORNE).

– Dernier paru, l'*Auzou débutant* (2008), qui contient d'assez nombreuses citations de contes et de fables, évoque l'appartenance de certains mots à cet univers textuel en recourant tour à tour à la définition (« Dans les histoires, une fée est une femme gentille qui a des pouvoirs magiques. » s.v. FÉE), à un exemple (« Dans les légendes, la licorne représente la pureté. » s.v. LICORNE), à une remarque (« Les revenants existent seulement dans les histoires. » s.v. REVENANT) ou à un développement encyclopédique (« Les ogres existent seulement dans les contes de fées. » s.v. OGRE). Il fournit aussides indications diaphasiques, usuellement dans des définitions (« Dans le langage familier, un copain, une copine est un(e) ami(e). » s.v. COPAIN, COPINE) et exceptionnellement dans une remarque (« Le mot *pagaille* est un mot familier. » s.v. PAGAILLE), sans d'ailleurs toujours marquer des items que d'autres répertoires marquent quand ils les sélectionnent (par exemple DINGUE, GOSSE, INFOS, MAGNANIME ou MOCHE). Enfin, on y rencontre de manière épisodique une appréciation diafréquentielle (« On peut aussi écrire *faignant*, *faignante*, mais c'est plus rare. » s.v. FAINÉANT, FAINÉANTE).

Tous les autres ouvrages semblent pouvoir être passés sous silence.

CONCLUSION

Avec cette contribution de format modeste, nous avons voulu construire un socle pour de possibles analyses plus fouillées, en décrivant quelques grandes lignes de l'évolution de la lexicographie française depuis trois décennies en matière de traitement de la variété des usages lexicaux, qui ne se limite pas à l'emploi de marques codifiées mais met à contribution d'autres composants des articles :

- progrès mesurés dans l'explicitation des dispositifs de marquage mis en œuvre dans les dictionnaires « encyclopédiques » généraux,
- tendance d'ensemble au délestage du diastratique au bénéfice du diaphasique,
- place accrue des francophonismes dans la variation diatopique,
- apparition de répertoires pour les cycles 2 et 3 de l'école primaire qui, avec leurs moyens propres, font aux faits de variation une place proportionnée à leurs finalités spécifiques.

Il resterait beaucoup à dire, et d'autres publications s'y emploieront peut-être, tant pour entrer plus en détail dans la description et l'évaluation des pratiques d'ouvrages seulement évoqués ici que pour analyser les évolutions observables dans des classes de répertoires spécialisés – dictionnaires de synonymes et antonymes ou de variétés régionales ou sociales de français – qui

n'ont pu être pris en compte et dont certains présentent des caractéristiques remarquables.

Il reste aussi beaucoup à faire pour les lexicographes sur les deux terrains critiques pointés par Corbin (1989, § 4) :

- la cohérence des traitements : il est déroutant, à trente ans d'intervalle et avec l'outillage informatique aujourd'hui disponible, que le *Dictionnaire du français* de 1999 soit, de ce point de vue, aussi aisé à prendre en défaut que le *Micro Robert* de 1971, dont D. & P. Corbin (1980) avaient mis en évidence le déficit de rigueur (v. *supra* § 2.1) ;
- l'adéquation de la description des usages, qui, en dépit des inflexions observées, souffre toujours du carcan athéorique et idéologisé de marques réductrices insuffisamment articulées et reste biaisée par le poids endémique de la normativité, même dans les répertoires les plus ouverts aux usages « non standard » (v. Corbin 2006 : 45-47 et 111-114).

Comme naguère, le premier déficit, qui ne requiert que volonté et moyens, est en principe plus aisé à combler que le second, qui suppose l'élaboration d'un modèle de la variation des usages lexicaux pouvant fournir une alternative valide aux représentations sociolinguistiques préconstruites toujours en vigueur.

RÉFÉRENCES BIBLIOGRAPHIQUES

RÉPERTOIRES

Auzou débutant = *Dictionnaire Auzou débutant*, Paris, Éditions Philippe Auzou, 2008.

Auzou junior = *Dictionnaire Auzou junior*, Paris, Éditions Philippe Auzou, 2006.

CELLARD Jacques et REY Alain, 1980, *Dictionnaire du français non conventionnel*, Paris, Hachette.

Dictionnaire culturel en langue française, 4 vol., Paris, Dictionnaires Le Robert, 2005.

Dictionnaire de l'Académie française, 9^e éd., tome I, Paris, Imprimerie Nationale, 1992.

Dictionnaire du français langue étrangère. Niveau 1, Paris, Larousse, 1978 ; Niveau 2, 1979.

Dictionnaire du français. Référence. Apprentissage, Paris, Dictionnaires Le Robert et CLE International, 1999.

Dictionnaire encyclopédique Auzou, Paris, Éditions Philippe Auzou, 2002.

Dictionnaire général = *Dictionnaire général pour la maîtrise de la langue française, la culture classique et contemporaine*, Paris, Larousse, 1993.

Dictionnaire Hachette encyclopédique, millésime 1994, Paris, Hachette, 1993 ; millésime 2002, 2001.

DUPRÉ Paul, 1972, *Encyclopédie du bon français dans l'usage contemporain*, 3 vol., Paris, Éditions de Trévise.

Fleurus junior = *Fleurus junior. Dictionnaire encyclopédique*, Paris, Éditions Fleurus, 2001.

Grand Dictionnaire de français : voir *Mon grand dictionnaire illustré*.

Grand Larousse = *Le Grand Larousse illustré. Dictionnaire encyclopédique en 3 volumes et 1 CD-Rom*, Paris, Larousse, 2005.

Grand Robert = *Le Grand Robert de la langue française. Dictionnaire alphabétique et analogique de la langue française*, 2^e éd. entièrement revue et enrichie, 9 vol., Paris, Dictionnaires Le Robert, 1985.

Hachette benjamin = *Dictionnaire Hachette benjamin*, Paris, Hachette, 1996.

Hachette junior(s) = *Dictionnaire Hachette juniors*, Paris, Hachette, 1980, rééd. 1986 ; *Dictionnaire juniors 1. Langue française*, 1990 ; *Dictionnaire Hachette juniors*, 1993 ; *Dictionnaire Hachette junior*, 1998.

Larousse des débutants : voir *Mini débutants*.

Larousse du collège, Paris, Larousse, 2003.

Larousse junior, Paris, Larousse, 2003 ; nouv. éd. *Dictionnaire junior*, 2008.

Larousse pratique. Dictionnaire du français au quotidien, Paris, Larousse, 2003.

Logos. Grand dictionnaire de la langue française, 3 vol., Paris, Bordas, 1976.

Maxi débutants : voir *Nouveau Larousse des débutants*.

Maxidico = *Dictionnaire encyclopédique de la langue française. Le Maxidico*, millésime 1997, [Bruxelles,] Éditions de la Connaissance, 1996.

Micro-Robert. Dictionnaire du français primordial, Paris, SNL - Le Robert, 1971 ; nouv. éd. : *Le Micro-Robert. Dictionnaire d'apprentissage de la langue française*, Dictionnaires Le Robert, 1988.

Mini débutants = *Mini débutants. Mon premier vrai dictionnaire*, Paris, Larousse, 1985 ; nouv. éds *Dictionnaire Mini débutants*, 1999, *Larousse des débutants*, 2005.

Mon grand dictionnaire illustré, Paris, Éditions M et L et Éditions Philippe Auzou, 2000 ; nouv. éd. *Le Grand Dictionnaire de français*, [Mantes-la-Jolie,] Éclairs de plume, 2004.

Nouveau Larousse des débutants, Paris, Larousse, 1977 ; nouv. éds *Maxi débutants. Le dictionnaire CE2, CM*, Larousse, 1986 ; *Dictionnaire Maxi débutants*, 1997.

Nouveau Petit Robert : voir *Petit Robert*.

Petit Fleurus = *Le Petit Fleurus*, Paris, Éditions Fleurus, 1998.

Petit Larousse = *Nouveau Petit Larousse illustré*, millésime 1938, Paris, Larousse, 1937 ; autres éds : *Petit Larousse*, millésime 1960, 1959 ; *Petit Larousse illustré*, millésime 1981, 1980 ; millésime 1989, 1988 ; *Le Petit Larousse illustré*, millésime 1992, 1991 ; millésime 1998, 1997 ; millésime 2000, 1999 ; *Le Petit Larousse*, millésime 2003, 2002 ; *Le Petit Larousse illustré*, millésime 2005, 2004.

Petit Robert = *Le Petit Robert. Dictionnaire alphabétique et analogique de la langue française*, Paris, SNL, 1967 ; nouv. éd. *Le Nouveau Petit Robert. Dictionnaire alphabétique et analogique de la langue française*, Paris, Dictionnaires Le Robert, 1993.

Petit Robert des enfants = *Le Petit Robert des enfants. Dictionnaire de la langue française*, Paris, Dictionnaires Le Robert, 1988.

PICOCHÉ Jacqueline et ROLLAND Jean-Claude, 2002, *Dictionnaire du français usuel*, Bruxelles, De Boeck-Duculot.

RÉZEAU Pierre (éd.), 2001, *Dictionnaire des régionalismes de France. Géographie et histoire d'un patrimoine linguistique*, Bruxelles, De Boeck-Duculot.

Robert benjamin = *Le Robert benjamin*, Paris, Dictionnaires Le Robert, 1997.

Robert brio : voir *Robert méthodique*.

Robert collègue = *Le Robert collègue*, Paris, Dictionnaires Le Robert, 1997.

Robert de poche = *Le Robert de poche*, Paris, Dictionnaires Le Robert, 1995 ; nouv. éd. 2006.

Robert junior = *Le Robert junior illustré*, Paris, Dictionnaires Le Robert, 1993, nouv. éd. 1999, 2005.

Robert méthodique = *Le Robert méthodique. Dictionnaire méthodique du français actuel*, Paris, Dictionnaires Le Robert, 1982 ; nouv. éd. *Le Robert brio. Analyse des mots et régularités du lexique*, Paris, Dictionnaires Le Robert, 2004.

Robert pour tous = *Le Robert pour tous. Dictionnaire de la langue française*, Paris, Dictionnaires Le Robert, 1994.

Robert quotidien = *Le Robert quotidien. Dictionnaire pratique de la langue française*, Paris, Dictionnaires Le Robert, 1996.

Scolaire Hachette = *Dictionnaire scolaire Hachette*, Paris, Hachette, 2002.

Super major = *Dictionnaire super major*, Paris, Larousse, 1994 ; nouv. éd. 1997, *Larousse super major*, 2004.

Trésor de la langue française. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789-1960), 16 vol., Paris, Éditions du CNRS (t. 1-10) et Gallimard (t. 11-16), 1971-1994.

AUTRES RÉFÉRENCES

CORBIN Danielle et CORBIN Pierre, 1980, « Le monde étrange des dictionnaires (1). Les “marques d’usage” dans le *Micro-Robert* », *Bulletin du Centre d’analyse du discours* 4, Lille, Presses Universitaires de Lille, p. 237-324.

CORBIN Pierre, 1989, « Les marques stylistiques / diastratiques dans le dictionnaire monolingue », in F.J. Hausmann, O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter, I, p. 673-680.

CORBIN Pierre, 2006, *Avec des dictionnaires pour compagnons*, habilitation à diriger des recherches, vol. 1, Université Charles de Gaulle - Lille 3 / L’Atelier C.

GADET Françoise, 2003, *La Variation sociale en français*, Gap et Paris, Ophrys.

GASIGLIA Nathalie, 2008a, « Le traitement des emprunts dans les dictionnaires d’apprentissage français : options descriptives et choix rédactionnels », dans J. Pruvost (éd.), *Les Journées des dictionnaires de Cergy. Dictionnaires et mots voyageurs. Les quarante ans du « Petit Robert »*, Éragny-sur-Oise, Éditions des Silves, collection « Actes de colloque », p. 153-212.

GASIGLIA Nathalie, 2008b, “Description of loan words in French school dictionaries: treatment of words of foreign origin in *Dictionnaire Hachette junior* (2006) and *Le Robert junior illustré* (2005)”, in E. Bernal and J. DeCesaris (eds.), *Proceedings of the XIII EURALEX International Congress (Barcelona, 15-19 July 2008)*, livre + CD-Rom, Barcelona, Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada, p. 1115-1122.

- GASIGLIA Nathalie, 2008c, « Le traitement des anglicismes dans quelques dictionnaires français pour jeunes lecteurs », dans F. Maniez et P. Dury (éds), *Lexicographie et terminologie : histoire de mots. Hommage à Henri Béjoint*, Gap, Louis Jean Imprimeur, collection « Travaux du CRTT », p. 157-174.
- HAUSMANN Franz Josef, 1977, *Einführung in die Benutzung der neufranzösischen Wörterbücher*, Tübingen, Max Niemeyer Verlag, collection « Romanistische Arbeitshefte », n° 19.
- HAUSMANN Franz Josef, 1989, „Die Markierung im allgemeinen einsprachigen Wörterbuch: eine Übersicht“, in F.J. Hausmann, O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter, I, p. 649-657.
- IMBS Paul, 1971, « Préface », *Trésor de la langue française I*, p. ix-xxvij.
- PICOCHÉ Jacqueline, 2002, « Présentation et mode d'emploi », *Dictionnaire usuel du français*, p. 7-26.
- REY Alain, 1967, « Présentation du dictionnaire », *Le Petit Robert*, p. ix-xxj.
- REY Alain, 1983, « Norme et dictionnaires (domaine du français) », dans É. Bédard et J. Maurais (éds), *La Norme linguistique*, Gouvernement du Québec, Conseil de la langue française et Paris, Le Robert, collection « L'ordre des mots », p. 541-569.
- REY Alain, 1985, « Préface de la deuxième édition », *Le Grand Robert de la langue française*, p. xvij-xlij.
- REY Alain, 2005, « Avant-propos, et après-faire », *Dictionnaire culturel en langue française*, I, p. xj-xxj.
- REY-DEBOVE Josette, 1988, « Préface », *Le Petit Robert des enfants*, p. ix-xviii.
- REY-DEBOVE Josette, 1999, « Dictionnaire du français. Avant-propos », *Dictionnaire du français*, p. vij-xj.
- REY-DEBOVE Josette et REY Alain, 1993, « Préface du Nouveau Petit Robert », *Le Nouveau Petit Robert*, p. ix-xxj.

LE POIDS DES CONTRAINTES DICTIONNAIRIQUES SUR L'ÉVOLUTION DES MARQUEURS DANS LES *PETIT LAROUSSE* (1997-2007)

Camille MARTINEZ
Université de Cergy-Pontoise

Résumé — Cet article s'appuie sur une comparaison des *Petit Larousse* millésimés 1997 à 2007. Nous décrivons tout d'abord le processus de mise à jour annuelle de ce dictionnaire : informations ajoutées, supprimées et modifiées. Après avoir établi que, pour des raisons économiques, seul un nombre restreint de pages est modifié lors d'une mise à jour, nous détaillons les choix lexicographiques et éditoriaux des auteurs, bridés par des contraintes dictionnaires. Nous constatons ensuite que les marqueurs lexicographiques de tous types sont largement touchés lors des mises à jour. Des marqueurs sont souvent ajoutés ou supprimés, pour gagner un peu de place ou au contraire pour apporter une information supplémentaire sans augmenter le nombre de lignes. Enfin, nous suivons l'évolution du seul marqueur *littéraire* dans la lettre E des *Petit Larousse* 1997 à 2007. Ce parcours dans la décennie est éclairé par les enseignements dégagés dans la première partie. C'est en somme de la caducité des marqueurs dont il est question, à la lumière des contraintes dictionnaires rarement décrites qui sévissent dans la fabrication d'un dictionnaire de référence.

Mots clés — évolution des nomenclatures, disparition des marques, contraintes dictionnaires, *Petit Larousse*.

Abstract — This article examines and compares several editions of the *Petit Larousse* (from 1997 till 2007). First the author describes how lexicographers update a dictionary yearly, and then gives examples of added, suppressed and modified information. The role of economic constraints is considered in relation to such choices. Finally, lexicographic and editorial choices made by the *Petit Larousse* are examined in great detail, where we find that almost all lexicographic labels are modified during the update, and the only reason they are left out is to make space for additional information without altering the total number of pages. The label *littéraire* in the letter E is used as an example.

Keywords — change in dictionary entries, change in labels, constraints in lexicography, *Petit Larousse*.

INTRODUCTION

« Un mot qui porte un marqueur dans le *Petit Larousse* est catégorisé de manière fiable. » « Les marqueurs du *Petit Larousse* sont mis à jour en fonction de l'évolution de la langue. » Telles sont les deux hypothèses que nous nous proposons de réfuter.

Il s'agit tout d'abord de démontrer que les apparitions et les disparitions de marqueurs dans les *Petit Larousse* ne sont pas nécessairement motivées par des raisons linguistiques, mais plus souvent par des contraintes d'ordre dictionnaire (Quemada 1987, Pruvost 2006 : 100-104) aux effets aléatoires : la plupart des évolutions de marqueurs se produisent lors d'ajouts d'articles nouveaux à proximité. De nombreux exemples viendront appuyer cette hypothèse, notamment une recherche ciblée sur la seule marque *littéraire* dans une tranche du *Petit Larousse*, la lettre E, durant la décennie 1997-2007. Nous nous interrogerons ensuite sur l'existence de raisons intrinsèques qui justifieraient les changements liés aux marqueurs, avant de conclure sur les rebondissements méthodologiques auxquels cet article conduit.

Nous appellerons ici « marque » un signe susceptible de catégoriser un mot, et « marqueur » la manifestation physique d'une marque dans un article de dictionnaire.

1. LE CARACTÈRE ALÉATOIRE DE L'INSERTION ET DE LA SUPPRESSION DES MARQUEURS

1.1 CONTRAINTES D'ORDRE DICTIONNAIRIQUE

Partons d'une constatation qui nous est apparue pendant la lecture comparée des *Petit Larousse* millésimés 1997 à 2007¹ : lors de la mise à jour annuelle du dictionnaire (mais non lors d'une refonte), les lexicographes ne modifient pas toutes les pages, mais seulement une sur trois environ dans la partie « langue française », tandis que les pages restantes sont strictement identiques d'une année à l'autre. Cela est lié à des contraintes économiques. En effet, la maison Larousse fait appel à un compositeur privé pour la confection du *Petit Larousse*, le compositeur étant une entreprise chargée d'assembler les caractères qui en constituent chaque page. Étant donné que le *Petit Larousse* est édité chaque année, il est demandé à cette entreprise de conserver les épreuves du dernier dictionnaire paru, en vue d'une réutilisation ultérieure. Ainsi, lorsqu'un nouveau *Petit Larousse* est fabriqué, il est inutile de le recomposer dans son intégralité, ce qui entraînerait des coûts élevés. Le compositeur ne modifie que les pages que les lexicographes et les éditeurs lui notifient, et laisse intactes celles, majoritaires, que les lexicographes ne veulent pas modifier. Cette astuce diminue le coût de production du nouveau dictionnaire, mais elle se mue en contrainte d'ordre dictionnaire, et influence directement le travail des lexicographes.

En effet, puisque le nombre de pages retouchées est restreint, il faut que la raison invoquée pour justifier cette retouche apparaisse indispensable. Un ajout (d'un nouveau mot ou d'un nouveau sens à un mot, opérations essentielles à la

¹ La comparaison de ces onze éditions successives du *Petit Larousse* avait pour cadre une recherche de thèse portant sur l'évolution de l'orthographe dans ces dictionnaires. Le présent développement sur les marques, qui profite des avancées méthodologiques de notre recherche, est une annexe de cette thèse.

mise à jour du *Petit Larousse*), ou une mise à jour liée à l'actualité, sont des raisons qui nécessitent à elles seules de modifier une page. En revanche, une modification mineure sur le texte (correction d'une erreur, harmonisation de deux articles semblables), ou bien une mise à jour reflétant une évolution linguistique, est rarement suffisante pour entraîner à elle seule la modification d'une page.

Le fait que, lors de la mise à jour d'un dictionnaire, des informations soient supprimées dans le but d'en ajouter de nouvelles, a déjà été évoqué dans divers ouvrages, notamment Corbin et Corbin (1989 : 941), Rey (1990 : 1826), Catach (1997 : 109), C. Muller (1999 : 82), Pivot (2004 : 13), R. Muller (2006 : 107), mais à notre connaissance aucune étude ne détaille pleinement le phénomène.

1.2 GAIN DE PLACE SUR UNE PAGE DU DICTIONNAIRE

Intéressons-nous maintenant au procédé d'insertion d'un nouvel article sur une page. Dans la plupart des cas de figure, une page modifiée est précédée et suivie par deux pages non modifiées. Cela implique que la première et la dernière ligne de la page retouchée doivent nécessairement être identiques d'une année à l'autre. Autrement dit, les lexicographes ne disposent d'aucune place prédéfinie pour insérer un nouvel article. Le nombre de lignes ajoutées sur la page doit donc au préalable être supprimé ailleurs. Pour cela, trois solutions : la moins utilisée est la coupure d'une illustration. Par exemple, lorsque le mot ARABO-ANDALOU entre dans le *Petit Larousse* 2006, l'illustration d'ARABESQUE est lésée de quelques millimètres. La deuxième solution est la diminution de la justification verticale, c'est-à-dire de l'espacement qui sépare les lignes deux à deux. Par ce procédé presque invisible (la différence s'exprime en fractions de millimètre), une colonne de quatre-vingt-huit lignes peut être tassée pour en accueillir quelques-unes supplémentaires. Enfin, la solution la moins raisonnable mais la plus fréquente est la suppression ou la réécriture d'une partie du texte. Pour illustrer ces deux dernières opérations, observons deux cas réels d'ajout d'article dans le *Petit Larousse*.

Dans le *Petit Larousse* 2001, la page 132 accueille trois nouveaux mots : BEUGLANT, BIBANDE et l'article à double entrée BIBITE OU BÉBITE. Concrètement, ces trois articles représentent treize lignes à insérer dans la page. La stratégie utilisée pour gagner ces lignes est la diminution de la justification verticale. Le nombre de lignes total des trois colonnes passe de 252 en 2000 à 265 en 2001. Cependant, sur cette même page 132, cinq articles sont le théâtre de phénomènes annexes dont le but ne semble pas être de gagner de la place. Ainsi l'article à double entrée BICHLAMAR OU BICHELAMAR perd-il sa deuxième variante graphique pour se muer en entrée simple : BICHLAMAR. De plus, l'adjectif BEUR, *adj. inv. en genre* en 2000 devient simple *adj.* en 2001. Dans le même temps, la quatrième définition de BIAIS est remaniée et clarifiée. Enfin, la présentation de l'étymologie de BICHOF est un peu modifiée. Et enfin, et c'est le phénomène qui nous intéresse ici le plus, le marqueur de l'emploi transitif du verbe BIAISER (au sens de « Introduire un biais dans ») change de nature. De *sociol.*, il devient *stat.*

Dans le *Petit Larousse* 2002, le SOMONI fait son entrée page 950. Les deux lignes nécessaires à son insertion sont gagnées en supprimant l'adjectif SOM-MABLE. À l'article SONDE, le *trou de sonde* fait également son apparition, et la

ligne dont il a besoin pour s'implanter est créée en diminuant la justification verticale. Dans le même temps, sur la même page, les articles 3. SON, SOMESTHÉSIE, SOMPTUAIRE et SONDAGE sont légèrement réécrits dans le sens d'une réduction du texte, mais cela n'entraîne pas de gain de ligne. Et surtout, l'article SOMNIFÈRE perd son marqueur. La définition du *Petit Larousse* 2001, « Cour. Hypnotique », devient en 2002 simplement « Hypnotique ».

1.3 HIÉRARCHIE DES CHANGEMENTS SUR UNE PAGE DE DICTIONNAIRE

À travers ces deux exemples (parmi des milliers), nous venons de décrire plusieurs changements survenus sur des pages du *Petit Larousse*, lors de mises à jour. Le caractère ordonné de ces changements nous est apparu peu à peu, en huit mois de familiarité avec les *Petit Larousse* 1997 à 2007. Au fil de lectures comparées, une théorie s'est peu à peu construite. Cette théorie permet de reconstituer à rebours et de manière précise le travail des lexicographes soumis aux contraintes dictionnairiques que nous venons de décrire. L'idée de départ de cette théorie est fort simple : tous les changements constatés sur une page du *Petit Larousse* lors d'une mise à jour s'ordonnent hiérarchiquement. Nous proposons ici les principes de cette théorie, accompagnés d'une terminologie succincte et d'une application au niveau des marqueurs.

Au sommet de la pyramide se trouvent les *changements totalement nécessaires* (principalement des ajouts d'articles, de nouvelles définitions), qui font du *Petit Larousse*, chaque année, un nouveau dictionnaire. Ces changements sont comparables à des forces qui peuvent éventuellement entraîner différents types de conséquences, que nous organisons en trois groupes.

En premier lieu, des *conséquences directes et logiques*. Par exemple, lorsque l'article BIBITE OU BÉBITE entre sur la page 132 du *Petit Larousse* 2001, il commande l'entrée de l'article renvoi BÉBITE → BIBITE page 127. Remarquons qu'à son tour, sur la page 127, l'insertion de cet article renvoi apparaît comme un changement totalement nécessaire, pouvant entraîner de nouvelles conséquences.

En second lieu, des *conséquences indirectes et arbitraires*, liées à la recherche de place. Ainsi nous avons vu comment l'article SOMMABLE avait été supprimé pour laisser sa place vacante à l'article SOMONI. La suppression de SOMMABLE est arbitraire dans le sens où ce mot n'a aucun rapport avec le mot SOMONI, si ce n'est ses trois lettres initiales, qui ont déterminé sa place dans le dictionnaire.

Enfin, des *conséquences annexes*, qui ne sont pas en rapport direct avec le changement apporté sur la page, et qui ne permettent aucun gain de place. Elles comportent une part d'arbitraire dans le sens où elles ne surviennent que si une page est modifiée pour une autre raison et au niveau d'un autre article. Elles sont parfois motivées par des raisons liées à l'évolution de la langue (ajout d'une graphie, féminisation), ou à la recherche de perfection (rectification d'une anomalie, clarification d'une imprécision). Les conséquences annexes de la modification d'une page ont ceci de remarquable que, bien qu'elles contribuent à améliorer le dictionnaire, elles ne constituent pas une raison suffisante pour impliquer à elles seules la modification de cette page : elles ne sont pas rentables. Nous avançons donc l'idée qu'elles restent en sommeil plusieurs années avant

de survenir, opportunément, l'année où leur page connaît un changement totalement nécessaire, qui agit à la manière d'une catalyse.

Cette présentation de la hiérarchie des changements sur une page de dictionnaire doit permettre aux linguistes, aux métalexicographes et au public de reconstituer la raison qui a précipité, sur telle ou telle page, le changement (de marqueur ou autre) qu'ils ont constaté ou qu'ils étudient, à travers un questionnement simple qu'il nous reste à détailler. Au préalable, on aura identifié le changement, trouvé l'année à laquelle il est survenu, réuni les deux millésimes qui circonscrivent ce changement, et relevé toutes les modifications apportées à la page concernée.

1. Le dictionnaire dans lequel on observe un changement est-il le fruit d'une refonte, ou inaugure-t-il une nouvelle maquette ? Si oui, la théorie ne fonctionne pas, et le changement n'est probablement pas dû à des raisons économiques. Considérons le cas où le changement survient lors d'une mise à jour.
2. Le changement observé est-il seul sur la page ? Si oui, c'est qu'il est nécessaire et justifié. Sinon, sa nécessité est à établir.
3. Parmi toutes les modifications relevées sur la page, lesquelles paraissent les plus nécessaires (ajout d'article, ajout de définition, actualisation de données politiques ou scientifiques, correction d'une erreur conséquente, reformulation en propos politiquement corrects, etc.) ? Il peut également s'agir d'un renvoi, synonyme, antonyme ou autre associé à un article nouveau inséré sur une autre page (v. *supra* l'exemple de BIBITTE OU BÉBITE).
4. Si le changement observé ne semblait pas nécessaire, est-il semblable à un autre changement survenu sur une autre page ? Est-il dépendant d'un autre changement éloigné ? Si oui, il s'agit d'une *conséquence directe* de cet autre changement. (On se servira du jeu de renvoi éventuellement contenu dans l'article observé pour savoir quelles autres pages sonder.)
5. Si le changement observé ne semble pas être la conséquence directe d'un autre changement, a-t-il permis un gain de place ? Si oui, il s'agit de la *conséquence indirecte* d'un autre changement prioritaire sur la page. On peut dire de ce changement qu'il est contraint par des raisons qui ne sont pas linguistiques.
6. Enfin, si le changement observé ne correspond à aucun des cas de figure envisagés, c'est qu'il s'agit d'une *conséquence annexe* qui survient en marge d'autres modifications.

1.4 APPLICATION

Dans les deux exemples donnés *supra* § 1.2, nous avons rencontré des changements de marqueurs (passage de *sociol.* à *stat.* et suppression de *cour.*). En suivant le questionnement présenté au § 1.3, on établit que ces changements sont des conséquences annexes d'un ajout d'article sur leur page. On peut donc considérer que leur part d'arbitraire est moindre. Observons maintenant des exemples similaires en utilisant le questionnement, ce qui permet de partir des changements de marqueurs pour déterminer leur cause. Nous avançons que cette méthode d'investigation doit être mise en œuvre pour tout changement observé dans deux éditions différentes d'un même dictionnaire².

² Bien sûr, nous avons conscience que cette nouvelle méthode d'investigation est lourde et difficile à mettre en œuvre, mais elle nous semble véritablement essentielle, car elle permet de mesurer efficacement

Dans le *Petit Larousse* 2006, la deuxième définition d'ALCYON (« polype s'abritant dans une loge squelettique calcaire ») se voit enrichie du marqueur *zool.* Quant à l'adjectif ALDÉHYDIQUE, il perd le marqueur *chim. org.* qu'il portait en 2005. Sur la même page apparaît au même moment le *syndrome d'alcoolisme fœtal*, à l'article ALCOOLISME. Si l'ajout de ce dernier élément est fondamental dans le travail de mise à jour du dictionnaire et se présente comme le changement nécessaire ayant entraîné la modification de la page, on peut en revanche s'interroger sur la motivation du changement aux articles ALCYON et ALDÉHYDIQUE. Tout d'abord, mentionnons que l'ajout de quatre lignes à l'article ALCOOLISME doit être compensé ailleurs sur la page. Justement, en perdant son marqueur *chim. org.*, l'article ALDÉHYDIQUE passe de deux à une seule ligne. Il s'agit donc d'une conséquence indirecte et arbitraire de l'ajout du *syndrome d'alcoolisme fœtal*.

Reste à comprendre pourquoi un marqueur *zool.* est ajouté à l'article ALCYON. Cela vient probablement du fait que sa définition porte une trace d'ambiguïté. En effet, POLYPE, au début de la définition, est un mot polysémique qui peut renvoyer soit à un animal, soit à une tumeur. Les deux sens du mot POLYPE sont d'ailleurs définis dans le *Petit Larousse*, et ils portent chacun le marqueur approprié : *zool.* pour le premier, et *méd.* pour le second. On peut penser que c'est pour empêcher le lecteur de croire qu'un alcyon est une tumeur que les lexicographes ont attribué à ce mot le nouveau marqueur *zool.* Cet ajout de marqueur va donc dans le sens d'une précision, contrairement à la suppression du marqueur à l'article ALDÉHYDIQUE. En résumé, sur les deux changements de marqueurs observés sur la page, l'un n'était pas primordial et n'est survenu cette année-là que parce que sa page était retouchée (ALCYON), tandis que l'autre est survenu pour des raisons économiques (ALDÉHYDIQUE). Remarquons avec étonnement que l'un apporte de l'information pendant que l'autre en supprime.

Sur la page 1046 du même *Petit Larousse* 2006, le marqueur *bx-arts* est ajouté à la deuxième définition de TENTURE. L'article TENTE-ROULOTTE, inséré sur cette page à ce moment-là, n'y est probablement pas pour rien. Étant donné que l'ajout de marqueur à TENTURE ne sert pas à gagner quelques lignes, il s'agit certainement d'une conséquence annexe de l'arrivée de TENTE-ROULOTTE.

Les exemples sont encore nombreux, mais ceux que nous venons d'étudier suffisent pour comprendre le phénomène. Tentons à présent une approche quantitative en suivant l'évolution d'une seule marque dans une petite tranche du dictionnaire.

2. LE SORT DE LA MARQUE LITTÉRAIRE DANS LA LETTRE E DES PETIT LAROUSSE 1997 À 2007

D'après les auteurs du *Petit Larousse* 2007, un mot reçoit le marqueur *litt.* lorsqu'il est utilisé « surtout à l'écrit ou dans des relations sociales réglées par des conventions, et [qu'il] produit un effet de sérieux ou d'élégance ». Accorder le marqueur *litt.* à un mot semble être un jugement particulièrement subjectif ; et

le taux de hasard dans un changement. Étant donné que nous disposons déjà d'une base de données considérable sur les changements survenus dans les dictionnaires récents, nous proposons d'effectuer des investigations ponctuelles pour les personnes qui nous demanderaient de répondre en termes de « hiérarchie des changements » à leurs interrogations.

cela se ressent lorsque l'on compare un même extrait du *Petit Larousse* dans plusieurs de ses éditions successives.

Sur la ligne de départ, dans la lettre E³ du *Petit Larousse* 1997 (64 pages), se trouvent 170 articles portant un marqueur *littéraire*. Les cas de figure sont très nombreux : certains articles sont monosémiques, d'autres polysémiques ; certains portent un seul marqueur, d'autres plusieurs. Sur la ligne d'arrivée, la lettre E du *Petit Larousse* 2007 (63 pages), se retrouvent 197 articles marqués *litt.* Quel chemin ont emprunté tous ces articles marqués ? C'est ce que nous nous proposons de détailler.

Le millésime 1998 est une refonte du *Petit Larousse* 1997 ; le texte du dictionnaire est retravaillé dans son entier. On relève donc de nombreux changements dans l'utilisation de la marque *litt.* Deux articles marqués disparaissent (ÉPOUSÉE et ÉVOÉ), et un article marqué nouveau fait son entrée (ÉRUBESCENT). Trente-trois articles déjà présents en 1997 reçoivent un marqueur en 1998, parmi lesquels ENTRETENIR, ESCARCELLE, ÉTIOLEMENT, ERRATIQUE, ENIVREMENT. D'autres, polysémiques et déjà marqués *litt.* en 1997, obtiennent un marqueur supplémentaire : EFFACER et EMBRASEMENT notamment. À l'inverse, une vingtaine d'autres articles perdent le marqueur ou l'un des marqueurs qu'ils portaient en 1997 : ÉCHAPPÉE, ENTENDRE, ENSEVELIR ; et parfois ce marqueur est remplacé par un autre comme *sout.*⁴ (pour les mots ENFREINDRE, ENJOINDRE, EXCÉDER et EXTRACTION notamment), *fig.* ou encore *vx.* Enfin, le mot ÉDÉNIQUE connaît un destin particulier : en 1997, son unique définition « qui évoque le paradis terrestre » est marquée *litt.* ; en 1998 arrive un nouveau sens « qui procure un bonheur paradisiaque », et le marqueur, inexplicablement, démissionne de l'ancienne définition et s'installe devant la nouvelle.

Le millésime 1999 est une simple mise à jour du *Petit Larousse* 1998. Deux marqueurs *litt.* sont ajoutés à deux articles de la page 406, EURYTHMIE et EURYTHMIQUE. Cela n'est pas un hasard. La page 406 est justement celle qui héberge le mot EURO, un mot pour lequel les lexicographes devaient faire preuve d'une certaine souplesse autour des années 2000, car tout ce qui était rattaché à la monnaie unique européenne évoluait très rapidement⁵. En l'occurrence, la remarque du *Petit Larousse* 1998 à propos de la monnaie unique à l'article EURO (« La liste des États où elle entrera en vigueur à partir du 1^{er} janvier 1999 sera arrêtée... ») n'est plus vraie l'année suivante, date à laquelle la liste des États est connue, et d'ailleurs relayée par le dictionnaire. La page 406 est donc retouchée pour être au plus près de l'actualité, et l'ajout de marqueur à EURYTHMIE et EURYTHMIQUE profite de cette opportunité pour survenir.

3 La lettre E du dictionnaire, pas plus que n'importe quel échantillon limité à une tranche alphabétique, n'est représentative de l'ensemble du texte. Nous l'avons choisie parce que les marqueurs *litt.* y foisonnent. La seule démonstration tentée étant celle du caractère désordonné et aléatoire des changements de marqueurs, les données observées et les résultats déduits ne fluctuent pas en fonction du choix de l'échantillon. Les exemples développés dans la première partie sont d'ailleurs issus d'autres endroits du dictionnaire.

4 Abréviation de « soutenu », utilisé dans le *Petit Larousse* pour marquer qu'un « mot [est] employé dans des relations sociales réglées par des conventions et produisant un effet de sérieux ou d'élégance ».

5 Signalons pour l'anecdote que certaines pages du *Petit Larousse* sont ainsi plus souvent modifiées que d'autres, comme celle qui héberge l'article EURO, ou encore celle qui accueille le tableau listant les manifestations successives des Jeux olympiques.

Le *Petit Larousse* 2000 est une mise à jour du *Petit Larousse* 1999. Dans ce millésime, cinq articles marqués *litt.* sont totalement supprimés : ÉMERILLONNÉ, ENFIELLER, ÉPEURER, EXPÉDIENT et EXPIATEUR. La page où disparaît ÉPEURER est justement celle qui voit arriver l'article ÉPATAMMENT. Quant à la page 411 sur laquelle EXPÉDIENT et EXPIATEUR sont effacés, elle voit l'arrivée de l'adjectif EXPIABLE. Sur la page 380, c'est un nouveau sens à l'article ENFER qui provoque la disparition d'ENFIELLER. Quant à la page 373, qui voit la sortie d'ÉMERILLONNÉ, elle est retouchée pour des raisons que nous n'avons pas pu établir avec certitude. Quoi qu'il en soit, remarquons que les mots marqués *litt.* ont été des cibles privilégiées lors de cette mise à jour, lorsqu'il a fallu dégager de la place pour les nouveautés.

Le *Petit Larousse* 2001 est une mise à jour du *Petit Larousse* 2000. On voit s'y établir une erreur dans le marquage du verbe ÉRIGER. Dans les *Petit Larousse* 1997 à 2000, le marqueur *litt.*, placé avant la numérotation des différents sens, englobait l'ensemble de l'article. Dans le *Petit Larousse* 2001, elle est ajoutée – peut-être par erreur – au sens 2, si bien que l'on se retrouve dans un cas de figure que l'on peut schématiser ainsi : « *litt.* [1. (...) 2. *litt.* (...) 3. (...)] ». Ce marquage anormal n'a pas été modifié à ce jour et se retrouve tel quel dans le *Petit Larousse* 2007.

Le *Petit Larousse* 2002 est une version mise à jour du millésime 2001. Nous n'y avons pas relevé de changement dans l'utilisation de la marque *litt.*, du moins en ce qui concerne les mots commençant par la lettre E.

Le *Petit Larousse* 2003 est une mise à jour du *Petit Larousse* 2002. Nous y avons remarqué l'insertion d'un marqueur *litt.* page 404 à l'article ÉTONNER, marqueur attribué à un sens déjà présent dans les éditions antérieures. Constatons que cet ajout est concomitant avec l'arrivée d'un nouveau sens une ligne plus haut, dans le même article ÉTONNER. Remarquons au passage que cet ajout d'un nouveau sens, qui provoque la modification de la page en son entier, permet à d'autres marqueurs d'aller et venir ; ainsi le mot ÉTOUFFOIR devient-il *vieilli*, et les articles ÉTOUPE et ÉTRANGETÉ gagnent-ils en précision grâce aux marqueurs respectifs *text.* et *phys.*

Dans le *Petit Larousse* 2004, version mise à jour du *Petit Larousse* 2003, nous n'avons pas relevé de changement dans l'emploi de la marque *litt.*

Le *Petit Larousse* 2005 n'est ni une simple mise à jour, ni une refonte du millésime 2004, mais un compromis des deux. Il inaugure une nouvelle maquette, ce qui signifie que toutes les pages ont été modifiées. Toutefois, tous les articles n'ont pas été revus en profondeur, à l'exception des mots de certains domaines, comme l'informatique et l'automobile. Deux mots marqués littéraires ont disparu : ÉMERGEMENT et EFFICACE (le nom féminin, qui bénéficiait d'un article séparé de l'adjectif). Des marqueurs *litt.* sont apparus dans neuf articles déjà présents dans les éditions antérieures : ÉLECTION, ÉLASTICITÉ, ÉLECTIF, ENJOLIVEUR, ENJOLIVURE, ÉTANCHER, ERREMENTS, ESPÉRER et ÉTERNITÉ. À l'inverse, des marqueurs présents dans l'édition 2004 ont été supprimés en 2005, aux articles ÉGIDE et ÊTRES (son marqueur *litt.* est d'ailleurs remplacé par *vieux*). Enfin, signalons un fait intéressant : quatorze articles marqués *sout.* en 2004 deviennent *litt.* en 2005. Or parmi eux, quatre possédaient pourtant déjà ce marqueur *litt.* dans le *Petit Larousse* 1997 ! Il s'agit des mots ENFREINDRE, ENJOINDRE, EXCÉDER et EXTRACTION, que nous avons rencontrés plus haut. Les

dix autres articles ayant subi cette transformation de *sout.* à *litt.* en 2005 n'avaient reçu leur premier marqueur *sout.* que quelques années auparavant, dans le millésime 1998. Il s'agit des mots ÉBATTRE, ÉDIFIANT, ÉDIFIER, ÉDILE, EFFET, ENTACHER, ÉTROIT, EUPHORISER, EXCLUSIVE et EXEMPT.

Dans le *Petit Larousse* 2006, simple mise à jour du *Petit Larousse* 2005, nous n'avons noté aucun changement. Dans le *Petit Larousse* 2007, nous n'avons relevé qu'une modification, à l'article ÉGÉRIE. Au moment où cet article jusqu'alors monosémique se voit allongé de deux nouveaux sens, la définition d'origine du mot perd le marqueur *litt.* qu'elle portait depuis plusieurs années. Terminons sur une note fleurie en ajoutant que si la page 397, qui porte le mot ÉGÉRIE, est modifiée dans le millésime 2007, c'est non seulement pour ajouter deux sens à cet article, mais c'est aussi – espérons que c'est surtout ! – parce que l'illustration qui figure une vue des égouts dans leurs moindres détails est remplacée par le dessin, bien moins méphitique, d'une fleur d'églatier.

3. RECHERCHE D'UNE RAISON PROFONDE POUR LES CHANGEMENTS LIÉS AUX MARQUEURS

Nous avons observé comment des marqueurs étaient parfois supprimés pour gagner un peu de place sur une page où il était nécessaire d'en créer. Nous avons aussi décrit comment certains changements secondaires impliquant des marqueurs survenaient opportunément sur des pages, au moment précis où celles-ci étaient modifiées pour y ajouter de nouvelles informations essentielles dans le procédé de mise à jour du dictionnaire. Nous avons vu également que des marqueurs pouvaient être ajoutés, comme à l'article POLYPE, pour lever une ambiguïté. Mais quelle est la raison profonde qui motive de manière générale les ajouts et les changements de marqueurs ? Voici deux éléments de réponse.

Pour qu'un marqueur soit ajouté à un article préexistant alors que rien ne permettait de pressentir cet ajout, il faut que le marqueur soit sous-jacent à l'article, et donc que celui-ci possède un développement plus étendu ailleurs que dans le *Petit Larousse*. Cela peut être dans un dictionnaire plurivolumaire, ou dans une base de données.

3.1 INTERACTION ENTRE LE *PETIT LAROUSSE* ET LE *GRAND LAROUSSE*

Un dictionnaire plurivolumaire tel que le *Grand Larousse encyclopédique* (dix volumes parus de 1960 à 1964), par nature, contient plus d'informations qu'un *Petit Larousse*. Il est possible que ce dictionnaire, ainsi que tous les autres plurivolumaires Larousse, soit une source pour la rédaction des dictionnaires monovolumaires. On y trouve en effet les marqueurs *zool.* et *arts décor.* aux articles ALCYON et TENTURE, articles auxquels des marqueurs (*zool.* et *Bx-arts*) ont été ajoutés dans les *Petit Larousse*. Ces marqueurs ajoutés dans le millésime 2006 étaient donc déjà présents dans les années 1960 : ils n'ont pas été créés *ex nihilo*.

3.2. MARQUAGE SOUS-JACENT DES NÉOLOGISMES

La banque néologique du *Petit Larousse* est une base de données dans laquelle tous les néologismes observés par les lexicographes sont décrits. Une partie de cette description est le marquage systématique de chaque néologisme. Sur chaque fiche sont codés l'origine géographique du mot (aire francophone ou non), qui peut prendre 225 valeurs différentes et combinables ; le niveau de

langue du mot, facultatif, qui peut prendre 13 valeurs ; et la discipline concernée, ce champ étant rempli obligatoirement. Les 97 valeurs possibles dans le champ discipline sont ordonnées par groupes ; par exemple la section sciences de la vie est divisée en sous-sections biochimie, biologie, etc. Un multicodage est possible lorsque le terme que l'on décrit est commun à plusieurs disciplines. Enfin, si le néologisme est un mot courant qui ne correspond à aucune discipline, il est tout de même codé par un marquage spécial. Ces outils de description des néologismes sont conçus pour commander des requêtes précises, mais l'aspect qui nous intéresse dans le marquage par disciplines est que celui-ci est obligatoire, ce qui implique que tous les néologismes qui entrent chaque année dans le *Petit Larousse*, depuis que la base de données existe, portent des marqueurs en puissance, susceptibles d'être apparents dans le texte imprimé, de rester sous-jacents, ou encore de changer d'état (en cas de multicodage).

Cependant, même après avoir étudié pendant plus d'une année les onze *Petit Larousse* mentionnés, nous ne sommes pas en mesure d'expliquer avec certitude ce qui motive les allées et venues des marqueurs dans les *Petit Larousse* de la dernière décennie. En dernier recours, nous avancerons que l'aspect humain – plutôt que mécanique – du travail lexicographique est un facteur qui génère de l'incertain dans le produit fini qu'est le dictionnaire.

CONCLUSION

Les lexicographes et les éditeurs du *Petit Larousse* fabriquent un nouveau dictionnaire chaque année. Leur travail est limité par les contraintes économiques que nous avons évoquées, si bien qu'il leur est parfois impossible d'apporter les modifications qu'ils souhaitent, tandis que d'autres fois, ils sont forcés d'exécuter des modifications qu'ils ne souhaitent pas⁶. Lors des refontes, pendant lesquelles les contraintes économiques s'estompent, ils travaillent avec un plus grand degré de liberté ; c'est ce qui explique la multitude des changements que nous avons décrits autour de la marque *litt.* dans les *Petit Larousse* 1998 et 2005. Toutefois, sur l'ensemble de la décennie, c'est un sentiment de confusion générale qui se dégage. Il ressort en effet que la description linguistique que propose le *Petit Larousse* s'enrichit et s'appauvrit de façon simultanée, à chaque fois qu'un marqueur apparaît ou disparaît, et que les changements surviennent de façon à la fois ordonnée et aléatoire.

Le but de notre travail est d'avertir les utilisateurs sur les limites du dictionnaire : un dictionnaire n'est en aucun cas un Code civil de la langue. Nous venons de démontrer qu'il n'était par exemple pas en mesure de signaler avec rigueur quels sont les mots qui produisent ou non « un effet de sérieux ou d'élégance ». Il en est de même avec d'autres aspects du dictionnaire (prononciation, étymologie, orthographe), qui sont mal appréhendés parce qu'ils subissent, comme les marqueurs, les effets désastreux de certaines contraintes économiques.

Loin des théories et des recherches des linguistes sur les marques lexicographiques, les auteurs du *Petit Larousse* font parfois de leur dictionnaire un terrain d'essai, dans le plus grand secret. Le classement hiérarchique des changements sur une page de dictionnaire, ainsi que l'investigation par questionnement qui en

6 Ces deux affirmations proviennent du témoignage oral d'un lexicographe.

découle, doivent à présent permettre aux linguistes et aux métalexicographes de mesurer avec plus de précision qu'auparavant quelles sont les parts du hasard et du fondement, de la vérité linguistiques dans un changement observé.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Collectif, 1960-1964, *Grand Larousse encyclopédique*, Paris, Larousse.
- Collectif, 1996, *Le Petit Larousse illustré 1997*, Paris, Larousse.
- Collectif, 1997, *Le Petit Larousse grand format 1998*, Paris, Larousse – Bordas.
- Collectif, 1998, *Le Petit Larousse illustré 1999*, Paris, Larousse – Bordas.
- Collectif, 1999, *Le Petit Larousse grand format 2000*, Paris, Larousse/HER.
- Collectif, 2000, *Le Petit Larousse illustré 2001*, Paris, Larousse/HER.
- Collectif, 2001, *Le Petit Larousse grand format 2002*, Paris, Larousse/VUEF.
- Collectif, 2002, *Le Petit Larousse 2003*, Paris, Larousse/VUEF.
- Collectif, 2003, *Le Petit Larousse illustré 2004*, Paris, Larousse/VUEF.
- Collectif, 2004, *Le Petit Larousse illustré 2005*, Paris, Larousse.
- Collectif, 2005, *Le Petit Larousse illustré 2006*, Paris, Larousse.
- Collectif, 2006, *Le Petit Larousse illustré 2007*, Paris, Larousse.
- Collectif, 1997, *Le Petit Larousse 1.01*, cédérom PC, Paris, Larousse – Liris Interactive.
- Collectif, 1997, *Le Petit Larousse 1.02*, cédérom PC, Paris, Larousse – Havas Interactive.
- Collectif, 1999, *Le Petit Larousse 2000*, cédérom PC, Paris, Havas Interactive – Larousse/HER.
- Collectif, 2000, *Le Petit Larousse 2001*, cédérom PC, Paris, Larousse/HER.
- Collectif, 2001, *Le Petit Larousse 2002*, cédérom PC, Paris, Larousse/VUEF.
- Collectif, 2002, *Le Petit Larousse 2003*, cédérom PC, Paris, Larousse/VUEF.
- Collectif, 2003, *Le Petit Larousse 2004*, cédérom PC, Paris, Larousse/VUEF.
- Collectif, 2004, *Le Petit Larousse 2005*, cédérom PC, Paris, Larousse.
- Collectif, 2005, *Le Petit Larousse 2006*, cédérom PC, Paris, Larousse.
- Collectif, 2006, *Le Petit Larousse 2007*, cédérom PC, Paris, Larousse.
- BIEDERMANN-PASQUES Liselotte et JEJCIC Fabrice (éds), 2006, *Les Rectifications orthographiques de 1990. Analyse des pratiques réelles (Belgique, France, Québec, Suisse, 2002-2004)*, Orléans, Presses universitaires d'Orléans.
- CATACH Laurent, 1997, « L'informatisation des dictionnaires de langue "Le Robert" », dans J. Pruvost (éd.), *Les Dictionnaires de langue française et l'informatique. La Journée des dictionnaires 1995*, Cergy-Pontoise, CRTH / UCP.
- CORBIN Pierre, 1989, « Les marques stylistiques / diastratiques dans le dictionnaire monolingue », dans F.J. Hausmann, O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter, p. 673-680.

- CORBIN Pierre et Corbin Danielle, 1989, « Sélection et description des dérivés et composés dans le dictionnaire monolingue », dans F.J. Hausmann, O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter, p. 937-946.
- HAUSMANN Franz Joseph, REICHMANN Oskar, WIEGAND Herbert Ernst und ZGUSTA Ladislav (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter.
- HAUSMANN Franz Joseph, 1989, „Die Markierung im allgemeinen einsprachigen Wörterbuch: eine Übersicht“, in F.J. Hausmann, O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter, p. 649-657.
- MARTINEZ Camille, 1999, *L'Évolution de l'orthographe dans les Petit Larousse et les Petit Robert 1997-2008 : une approche généalogique du texte lexicographique*, thèse de l'Université de Cergy-Pontoise.
- MARTINEZ Camille, 2004, *Stage aux éditions Larousse : Alimentation de la banque néologique du Petit Larousse*, rapport de stage, Bibliothèque des sciences du langage de l'Université Stendhal - Grenoble 3.
- MULLER Charles, 1999, *Monsieur Duquesne et l'orthographe. Petite chronique française 1988-1998*, Paris, CILF.
- MULLER Romain, 2006, « Les nouveautés dans les correcteurs orthographiques. Le point sur les correcticiels rectifiés », dans L. Biedermann-Pasques et F. Jejcic (éds), *Les Rectifications orthographiques de 1990. Analyse des pratiques réelles (Belgique, France, Québec, Suisse, 2002-2004)*, Orléans, Presses universitaires d'Orléans, p. 105-110.
- PIVOT Bernard, 2004, *100 mots à sauver*, Paris, Albin Michel.
- PRUVOST Jean (éd.), 1997, *Les Dictionnaires de langue française et l'informatique. La Journée des dictionnaires 1995*, Cergy-Pontoise, CRTH/UCP.
- PRUVOST J., 2006, *Les Dictionnaires français outils d'une langue et d'une culture*, Paris, Ophrys.
- QUEMADA Bernard, 1987, « Notes sur lexicographie et dictionnairique », *Cahiers de lexicologie*, 51, p. 229-242.
- REY Alain, 1990, « La lexicographie française depuis Littré », dans F.J. Hausmann, O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter, p. 1818-1843.

LES MARQUES D'USAGE DANS LE DICTIONNAIRE DU GREC STANDARD DE L'INSTITUT D'ÉTUDES NÉOHELLÉNIQUES (FONDATION MANOLIS TRIANTAPHYLLIDIS)¹

Anna ANASTASSIADIS-SYMÉONIDIS
Université de Thessalonique

Résumé — Après un tour d'horizon sur les marques de divers types utilisées dans le *Dictionnaire du grec standard*, nous examinons les marques d'usage concernant la norme linguistique et sociolinguistique, comme *savant*, *officiel*, *populaire*, *familier*, *argotique*, *vulgaire*, et nous procédons à une comparaison de vingt entrées, puisées dans la lettre B, entre ce dictionnaire et le *Dictionnaire du grec moderne* de Babiniotis. Nous constatons que la diversité et l'incertitude règnent sur les catégories et les critères utilisés pour assigner aux entrées les marques d'usage. La lexicographie, surtout celle qui concerne les dictionnaires de la langue générale et les dictionnaires pédagogiques, pourrait grandement profiter du travail des linguistes et des métalexigraphes s'ils avaient travaillé de manière approfondie sur ces phénomènes en grande partie encore inconnus. Pour y arriver, il faut s'orienter de la théorie à l'application et procéder à une description plus détaillée des unités lexicales au niveau du marquage.

Mots clés — marques diaphasiques, marques diastratiques, norme, savant, hypertexte.

Abstract — After reviewing different label typologies used in the Dictionary of Standard Greek (DSG), this article examines more specifically certain sociolinguistic labels such as *learned form*, *formal*, *slang*, *vulgar*, etc. Comparing the presentations of 20 entries, found under the letter B, in both the DSG and the Dictionary of Modern Greek (Babiniotis), the author underlines how diverse and sometimes vague are not only the above labels but also the criteria used to assign these very labels. Therefore, lexicography, especially that which deals with general usage of language, could greatly benefit from using the analyses and the results of studies carried out by linguists and metalexigraphers. Hence, theoretical underpinnings should be considered in both the description of words and in the assignment of labels.

Keywords — social labels, geographical labels, Greek dictionaries, hypertext, theory and lexicography.

¹ Je remercie les lecteurs anonymes pour leurs remarques ainsi que Mme Lucile Gravanis pour la lecture attentive de mon texte.

INTRODUCTION

Il est bien connu que les marques d'usage, malgré leur utilité indéniable, constituent un des points faibles des dictionnaires de la langue générale. Dans la première partie de cet article, nous allons présenter une typologie des marques d'usage utilisées en lexicographie, dans la deuxième partie, nous allons examiner les marques d'usage concernant la norme sociolinguistique utilisée dans le *Dictionnaire du grec standard* (DGS)², et plus particulièrement les termes *officiel*, *savant*, *familier*, *argotique*, *vulgaire* et *populaire*. Enfin, dans la troisième partie, nous allons comparer vingt entrées puisées dans la lettre B des deux dictionnaires généraux de la langue grecque parus en 1998, celui de Babiniotis et celui de l'Institut d'Études Néohelléniques (Fondation Manolis Triantaphyllidis), dans le but, d'une part, de démontrer l'incohérence du système des marques d'usage utilisé dans les deux dictionnaires et, de l'autre, d'améliorer cette situation par des propositions concrètes.

1. GÉNÉRALITÉS

Les niveaux de langue permettent de tenir des discours synonymes malgré les différences sociales et les conditions d'énonciation variées, et les marques d'usage permettent d'illustrer cette variation linguistique. On attribue, en général, aux marques d'usage, qui concernent l'usage linguistique et pas le référent, deux fonctions : soit une fonction normative/prescriptive soit une fonction descriptive ; dans le premier cas, l'écart qu'elles marquent par rapport à la langue standard impose la condamnation du mot marqué, tandis que dans les marques à fonction descriptive nous voyons une façon de protéger l'utilisateur du dictionnaire contre les abus de langage³. Dans ce travail, nous n'allons nous occuper que de ces dernières. De même, il ne faut pas oublier que les marques d'usage s'inscrivent dans le temps, car elles suivent, elles aussi, l'évolution de la langue et de la société. Ainsi, en ce qui concerne le grec moderne (GM), des emplois de la langue familière ont fini par s'intégrer dans la norme après 1974. On trouve des traces de cette translation dans la macrostructure (plus spécialement la nomenclature) et la microstructure des dictionnaires (les marques d'usage).

Il est nécessaire de connaître les conditions d'emploi d'une unité lexicale dans la société pour le décodage mais avant tout pour l'encodage, rôle assumé surtout par le marquage. Toutefois, l'opération de marquage⁴ relève du discours du lexicographe sur le système linguistique des énonciateurs (Mazière 1990 : 93), puisque c'est au moyen des marques que sont transmis des jugements métalinguistiques concernant avant tout les éléments lexicaux. Plusieurs linguistes se sont penchés sur la question : Imbs (1971 : xxxii-xxxiv), Dubois & Dubois (1971 : 37), Zgusta (1971 : 172), Corbin (1989 : 673). Plus exactement, pour Rey (*Le Petit Robert* 1967, 1969), les marques d'usage précisent la valeur d'emploi, soit dans le temps (ex. *vieux*, *vieilli*), soit dans l'espace (ex. *régional*),

2 Pour plus de détails v. Anastassiadis-Syméonidis & Tzivanopoulou 1984.

3 Selon Imbs (1971 : xvj) « les précisions sur les conditions d'emploi sont aussi des limitations contraignantes à la liberté d'emploi ».

4 Il y a deux manières de présenter le marquage : (a) le marquage explicite opéré par des termes appartenant au métalangage lexicographique, et (b) le marquage opéré par la seule typographie comme les flèches, le point d'exclamation et d'autres symboles.

soit dans la société, soit dans la fréquence (ex. *rare*), soit enfin dans le style (ex. *didactique*). Les marques d'usage ayant trait à la société sont les suivantes : *familier*, « courant dans la langue parlée ordinaire et dans la langue écrite un peu libre », et *populaire*, « courant dans les milieux populaires des villes, mais réprouvé ou évité par l'ensemble de la bourgeoisie cultivée ».

Rey-Debove (1971 : 91 et 1982 : x-xj) distingue quatre types de variation du langage : la variation spaciale, sociale, domainiale et temporelle ; elle appelle *niveau de langue* « toute caractérisation à l'intérieur de chacun de ces types, qui constituera un cas marqué ou non marqué ». La mention du niveau de langue est effectuée par les marques et elle est obligatoire dans le cas marqué, leur absence signifiant que le mot « est d'un emploi courant pour toute personne en toute situation et en tout lieu » (Rey-Debove 1982 : xj). Dans le cas de la variation sociale, elle reconnaît *populaire*, « qui qualifie un mot dans la langue parlée des milieux populaires », *argotique*, « mot réservé à un milieu particulier, surtout professionnel, mais inconnu du grand public », *familier*, « mot usité dans la langue parlée et écrite courante mais que l'on n'emploie pas dans des circonstances solennelles », *trivial* ou *vulgaire*, *langage soutenu* ou *académique*, *enfantin*, « mot du langage des jeunes enfants que peuvent employer les adultes en leur parlant ».

Coseriu (1976 : 13) propose une distinction tripartite des différences que présente une langue : il s'agit des différences diatopiques, diastratiques et diaphasiques. Les premières concernent l'espace géographique et ont trait aux dialectes, les secondes concernent les couches socio-culturelles et ont trait aux niveaux de langue, enfin les troisièmes concernent les types « situationnels » de modalité expressive et ont trait aux styles de langue.

Hausmann, dans son macromodèle du marquage dans les dictionnaires (1989 : 651), utilise 11 critères pour classer les marques :

- La temporalité, marquage diachronique, qui utilise des marques comme *vieux*, *vieilli*, *néologisme* ;
- La spacialité, marquage diatopique, qui utilise des marques comme *dialectal*, *régional*, concerne les indications d'ordre géographique, et marque l'appartenance de l'entrée à telle ou telle région linguistique, donc la dispersion géographique ;
- La nationalité, marquage diaïntégratif, qui utilise des marques comme *anglicisme* ;
- La médialité, marquage diamédial, qui concerne le médium d'expression et utilise des marques comme *écrit*, *oral* ;
- Les groupes socioculturels, marquage diastratique, qui utilise des marques comme *familier*, *populaire* ;
- La formalité, marquage diaphasique, qui utilise des marques comme *officiel* ;
- Les genres, marquage diatextuel, qui concerne le type de discours et utilise des marques comme *administratif*, *poétique* ;
- La technicité, marquage diatechnique⁵, qui utilise des marques comme *botanique*, *viticulture*. Les marques diatechniques (domainiales) sont nombreuses

5 Par ex. en grec η μεγάλη πίεση [la grande tension] « haute tension » appartient à la norme dans la langue générale et son synonyme η συστολική πίεση « la tension systolique » constitue la norme dans la terminologie médicale, mais il prend la marque diaphasique [+savant] dans la langue générale.

et permettent le passage d'une définition générale à des acceptions particulières, comme *linguistique*, *archéologie*, etc. Elles sont liées à la levée d'homonymie et à la polysémie et signalent l'appartenance de l'entrée à un sous-ensemble lexical donné. Elles précèdent la définition, en italiques ou entre parenthèses ;

- La fréquence, marquage diafréquentiel, qui utilise des marques comme *rare* ;
- L'attitude, marquage diaévaluatif, qui utilise des marques comme *euphémisme*, *mélioratif*, *péjoratif*, *intensif* ;
- La normativité, marquage dianormatif, qui utilise des marques comme *incorrect*, *emploi critiqué*, qui concerne les écarts de langage et les infractions à l'usage général.

Enfin, Petit (2006 : 201) distingue des unités lexicales diaphasiquement marquées (*bagnole*), qui entrent en concurrence avec une unité lexicale classifiante (*voiture*). Il propose (*ibid.* : 214) que les unités lexicales non dotées d'un indicateur d'usage soient considérées comme des dénominations de droit et que l'absence de marque confère à ces unités par défaut le statut d'unités typiques. Le marquage, en revanche, présente la double particularité de renseigner sur les propriétés socio-culturelles d'une unité mais aussi sur son degré d'intégration au lexique (*vieilli*, *familier*, *vulgaire*, *populaire*, *soutenu*, *littéraire*). Petit (*ibid.* : 305, 433) souligne l'asymétrie sémiotique entre une unité lexicale du registre standard et une unité marquée, et précise (*ibid.* : 435) qu'une proportion non négligeable d'unités lexicales de registre familier sont dépourvues d'équivalent standard ou soutenu.

Parmi ce riche ensemble de marques, nous n'allons nous occuper que des marques diaphasiques et diastratiques, parce que ce sont les variables « conditions d'énonciation » et « société » qui constituent avant tout la source d'incohérence. Malgré les flottements qui caractérisent l'emploi de ces marques d'usage, nous allons essayer de délimiter le champ des jugements diaphasiques et celui des jugements diastratiques.

Un jugement diaphasique a trait aux registres stratifiant l'énonciation et concerne la stylistique et la pragmatique ; il caractérise l'usage (les marques d'usage : *officiel*, *familier*), certaines situations de communication (*par ironie*, *raillerie*, *injurieux*, etc.) surtout pour signaler les registres inférieurs à la norme (non marquée), rarement supérieurs à la norme (ex. *soutenu*, *savant*) (Haensch *et al.* 1982 : 385-389), pour repérer l'inscription lexicale des conditions d'énonciation.

Par contre, un jugement diastratique, à visée sociolinguistique, renvoyant à un type d'énoncé, généralement bas (ex. *populaire*, *trivial*, *vulgaire*), marque l'appartenance du locuteur à une couche sociale définie, il implique une caractérisation du locuteur par rapport à un groupe social auquel il appartient, ex. *argot*, sans prendre en compte la situation de communication, mais en pointant des usages socialement marqués. On inclut aussi dans les marques diastratiques les différences qui concernent le sexe et l'âge (Haensch *et al.* 1982 : 382-385).

2. LES MARQUES D'USAGE DANS LE DICTIONNAIRE DU GREC STANDARD (DGS)⁶

Pour parler de marques d'usage, qui, dans ce dictionnaire, sont descriptives et non normatives, il faut tout d'abord considérer le langage dans son fonctionnement (et non pas *in abstracto*) et aussi cerner la zone langagière décrite comme norme objective⁷, par rapport à laquelle une forme ou un sens linguistique est considéré divergent, déviant, exclu ou marginalisé.

Si l'on compare le *Dictionnaire de Proïas* (1933) avec le DGS (1998), on note non seulement des différences dues à l'évolution linguistique, mais avant tout une translation de la norme, puisque, dans le *Dictionnaire de Proïas*, selon l'idéologie dominante de l'époque, la *katharevousa*⁸ occupe la place de la langue standard écrite, par rapport à laquelle tout écart est marqué⁹.

En 1977, nous avons établi le système de marquage qui serait utilisé dans le *Dictionnaire du Grec Standard*, inspiré du *Petit Robert*. Là se trouve d'ailleurs le lien de cette communication avec la langue française. Par ailleurs, pour restreindre la subjectivité des lexicographes, qui étaient au nombre de sept et d'âges variés, nous les avons incités, étant la rédactrice en chef du projet, à n'affecter la marque à une unité lexicale qu'après concertation avec les autres lexicographes. Les marques d'usage font de ce dictionnaire un ouvrage qui répond, entre autres, aux besoins spécifiques des usagers étrangers de niveau avancé. Les abréviations des marques d'usage, placées entre parenthèses, sont présentes dans un grand nombre d'articles¹⁰ et précèdent les définitions ou les formes grammaticales¹¹. De même, parfois une unité est accompagnée de deux marques (*populaire* et *littéraire*). Voici les marques qui nous intéresseront ici, accompagnées de leurs conditions d'emploi (DGS 1998 : ιη') :

- Επίσημο (soutenu / formel) : se dit des mots, souvent savants, fréquents dans un discours appartenant à la langue administrative, religieuse, militaire, ou dans des situations de communication exigeant des échanges officiels ou hiérarchiques, comme une fête nationale, un discours politique¹².
- Λόγιο (savant) : se dit des mots qui ont un équivalent en langue générale, mais qui proviennent de la langue pure ou du grec ancien quant à leur formation ou emploi (v. aussi Anastassiadis-Syméonidis & Fliatouras 2005). Le terme le plus proche en français serait *savant*, qui est défini ainsi dans le tableau des abréviations du *Petit Robert* : « pas connu, employé ou compris par

6 En dépit de son titre, ce dictionnaire ne contient pas que les mots appartenant au registre standard.

7 Le concept de norme tel qu'il est employé ici n'a rien à voir avec l'idéologie de la langue ni avec la correction linguistique.

8 *Katharevousa* «langue pure». Il s'agit de la langue forgée par les puristes au XVIII^e et XIX^e s. sur le modèle du grec ancien.

9 Ceci a des incidences aussi bien au niveau de la macrostructure (choix et forme des entrées) que de la microstructure de ce dictionnaire (forme du métalangage) (v. Anastassiadis-Syméonidis 1999).

10 Le DGS utilise bien d'autres marques : επιστημονικό (scientifique), λογοτεχνικό (littéraire), παιδικό (enfantin), παρωχημένο (vieilli), προφορικό (oral), ειρωνικό (ironique), μειωτικό (péjoratif), συναισθηματικό (affectif), υβριστικό (injurieux), σκωπτικό (par moquerie), χλευαστικό (par raillerie).

11 Les marques concernent non seulement des unités lexicales mais aussi des sens et des formes grammaticales, ex. αφήνω «laisser» – άφησα– (λάτκóτροπο) άφηκα et αφήκα «j'ai laissé».

12 Ce terme n'a pas d'équivalent dans le tableau des abréviations du *Petit Robert*.

l'ensemble des usagers cultivés ». Le concept en grec éclaire les dimensions étymologique et morphologique. Le terme de *didactique*, défini dans le tableau des abréviations du *Petit Robert* : « courant dans un traité, un cours », n'a pas d'équivalent en grec, mais recouvre les concepts de savant, soutenu, en langue écrite.

- Οικείο (familier) : se dit des mots utilisés en famille et dans des situations de communication amicales, qu'on évite d'employer quand on s'adresse à un inconnu ou à un supérieur dans une hiérarchie professionnelle ou autre ¹³.
- Λαϊκό (argotique) : se dit des mots souvent de la langue orale, utilisés par des gens vivant en marge de la société ou par des jeunes. La traduction par *argotique* de ce terme n'est pas heureuse, car elle peut prêter à confusion, puisque la définition du terme français dans le tableau des abréviations du *Petit Robert* est la suivante : « limité à un milieu particulier, surtout professionnel, mais inconnu du grand public ». La bonne traduction en grec du terme français serait συνθηματική γλώσσα. Le terme français se réfère à des milieux professionnels (jargon) mais le terme grec à des milieux sociaux.
- Χυδαίο (vulgaire) : se dit des mots grossiers, des mots tabous, liés à la sexualité ¹⁴. Cette marque concerne la bienséance.
- Λαϊκότροπο (populaire) : se dit des mots dialectaux couvrant une aire étendue qui peuvent être connus par les habitants des grandes agglomérations et qui sont utilisés dans la littérature. La traduction par *populaire* de ce terme n'est pas heureuse, car elle peut prêter à confusion, puisque la définition du terme français dans le tableau des abréviations du *Petit Robert* est la suivante : « courant dans la langue parlée des milieux populaires des villes, mais réprouvé ou évité par l'ensemble de la bourgeoisie cultivée (souvent argot ancien répandu) ». Le terme grec a des liens avec le folklore, d'où une couleur littéraire et méliorative, contrairement au terme français qui réfère aux classes sociales.

Comme dans le paragraphe suivant nous allons procéder à la comparaison des deux dictionnaires du grec moderne, dans le but d'évaluer la cohérence définitionnelle des marques d'usage et leur distribution, nous présentons le système de marquage du *Dictionnaire de Babiniotis* (DB) tel qu'il apparaît dans le chapitre « La structure du dictionnaire ». Le système au moyen duquel des entrées et des emplois sont caractérisés est tripartite et concerne trois dimensions :

- comment (style) et combien (fréquence) une unité est utilisée, p. ex. αρχαιο-πρεπές « archaïque », λόγιο « savant », λαϊκό « non savant », οικείο « familial », σπάνιο « rare », etc. ;
- où elle est utilisée (le type de communication linguistique où un sens est employé), p. ex. λογοτεχνία « littérature », διάλεκτοι « dialectes », αργκό « argot », etc. ;

¹³ Nous considérons ce terme équivalent au français *familier*, dont la définition dans le tableau des abréviations du *Petit Robert* est la suivante : « courant dans la langue parlée ordinaire et dans la langue écrite un peu libre, mais ne s'emploierait pas dans les circonstances solennelles ».

¹⁴ Nous considérons ce terme équivalent au français *vulgaire*, dont la définition dans le tableau des abréviations du *Petit Robert* est la suivante : « choquant, qu'on ne peut employer entre personnes bien élevées, quelle que soit leur classe sociale ».

- pourquoi elle est utilisée (commentaire du locuteur), p. ex. ειρωνική χρήση « ironique », σκωπτική « par moquerie », υβριστική « injurieux », μειωτική « péjoratif », εκφραστική « expressif ».

En particulier, le terme de λόγιο « savant » est utilisé pour qualifier des mots ou des formes conservant la morphologie de la *katharevoussa* ou en accord avec le style du discours des personnes cultivées. Le terme de λαϊκό « non savant » qualifie des mots ou des formes des couches sociales populaires, qui présentent un écart par rapport au grec standard au niveau morphologique ou phonétique. Le terme αρχαιοπρεπές « archaïque » réfère à des mots du grec ancien ou de la langue archaïsante et le terme λαϊκότερο « populaire » qualifie des mots ou des formes de la langue parlée des provinciaux. Enfin les mots tabous ou référant à la sexualité sont précédés du point d'exclamation, dont la fonction est ici équivalente au terme de vulgaire.

Trois remarques s'imposent :

- Puisque les termes αρχαιοπρεπές « archaïque » et λόγιο « savant » sont définis par des critères morphologiques, il doit y avoir un chevauchement dans leur distribution ; le premier a l'air d'être le comparatif ou le superlatif du second.
- Les termes grecs λαϊκό « non savant » et λαϊκότερο « populaire » sont constitués par le même adjectif avec la seule différence que le second se trouve au comparatif. Pourtant leur sens est très différent : le premier concerne le comment et le second les cas où un mot est utilisé ; c'est la raison pour laquelle nous avons proposé les traductions ci-dessus.
- En ce qui concerne le terme εκφραστικός « expressif », il n'est défini que dans le corps du dictionnaire avec des critères avant tout phonétiques.

Au terme de cette comparaison rapide, nous voyons que les deux dictionnaires, comparables en plusieurs points (monolingues, publiés en 1998, en gros de la même taille et s'adressant au même public), présentent différents types de dissymétrie : ils utilisent non seulement une liste différente de marques d'usage mais aussi une marque (λαϊκό) ayant le même signifiant mais un signifié différent relevant de variables différentes – diaphasique quand il s'oppose à *savant* (DB) ; diastatique quand il réfère à des milieux sociaux (DGS).

3. LES MARQUES D'USAGE ET LA NORME (SOCIO)LINGUISTIQUE

La conclusion qu'on peut tirer de cette confrontation sur échantillon restreint¹⁵ du marquage dans les deux dictionnaires est qu'il y a plusieurs divergences dues probablement à plusieurs facteurs : non homogénéité sociale, pas de définition rigoureuse des marques, qui se situent, comme nous venons de le dire par rapport à une norme, qu'aucun critère, d'ailleurs, ne saurait définir d'une manière précise. De même, selon Dubois (1971 : 24), la limite entre les divers niveaux de langue est souvent difficile à définir. D'ailleurs l'accumulation de marques d'usage sur un seul mot ou emploi peut bien désigner le malaise des lexicographes.

¹⁵ Nous n'avons sélectionné que des entrées où, à l'exception d'une seule, il y avait des différences au niveau du marquage.

Entrées	DGS	DB
Βραχίονας « main »	1b savant	2. Ø
Βρε « indignation forte »	3. (injurieux)	1. précède des expressions péjoratives
Βρίθω « être plein »	savant	Ø
Βρισίδι « bordée d'injures »	familier	Ø syn. savant υβρεολόγιο
Βρομόξυλο « raclée »	familier	expressif-familier
Βρομόπαιδο « sale enfant »	métaphore, péjoratif, injur.	expressif-familier
Βρομόσκυλο « canaille »	2. métaphore, injurieux	expressif-familier
Βρομύλος « malpropre »	savant	par raillerie
Βροντοφωνάζω « tonner de la voix »	métaphore	par synecdoque
Βροχηδόν « dru »	savant	archaïque
Βρόχι « lacs »	populaire	Ø
Βροχικά « bronchite »	familier	de tous les jours, courant
Βρόχος « nœud coulant »	1. vieilli (obsolète)	Ø
Βρυσομάνα « source abondante »	populaire	non savant
Βρώσιμος « mangeable »	savant	Ø
Βύζαγμα « allaitement »	familier	Ø
Βυζί « mamelle »	familier	Ø
Βύθος « torpeur »	populaire	3. métaphore
Βυσσοδομώ « tramer »	savant	savant, à sens négatif
Βωβός « muet »	savant	savant

Une bonne description lexicographique exige que les lexicographes signalent les écarts par rapport à la norme linguistique. Mais comment y parvenir ? Jusqu'à présent tous les lexicographes grecs ont eu recours à leur intuition linguistique. Pour quelles raisons ces inconséquences persistent-elles ? De notre point de vue, il y a quatre raisons :

1. Étant donné qu'on ne dispose pas de système de marquage rigoureux, fondé sur une théorie (socio)linguistique, les lexicographes se basent sur leur intuition et donc sur une appréciation subjective. Mais subjectif ne signifie pas aléatoire, comme le précise bien Corbin (1989 : 677), puisqu'il existe quand même des recoupements entre dictionnaires dans l'usage des marques.¹⁶
2. La marque constitue un jugement très relatif, car elle se situe toujours par rapport à une norme (un niveau neutralisé, le langage standard écrit courant), qui doit être définie de façon précise et empirique. Il n'y a pas de marque en soi, mais par rapport à un ensemble d'usages, qui sont pourtant arbitrairement définis. Par conséquent, la marque est un concept relatif. Cette relativité prend aussi un autre visage : on constate qu'on a besoin de degrés à l'intérieur d'une même catégorie (*très familier, moins familier que*). Si l'on représente ces marques d'usage sur un continuum avec la norme au milieu et

¹⁶ Ces recoupements s'expliqueraient en partie par le fait que les dictionnaires se copient entre eux.

quelques points remarquables à droite et à gauche, l'espace entre deux marques n'est pas vide, mais il comprend les propriétés nuancées.

3. Les marques dans un dictionnaire de langue¹⁷ décrivent à la fois le système langagier et sa mise en oeuvre par des usages, aboutissant à la diversité imprévisible des types de discours¹⁸.
4. Par ailleurs, cette intuition des lexicographes, pour être prise en considération, doit être collective et partagée par la majorité des locuteurs. De même les marques sont en étroite relation avec l'objet à décrire, tel qu'il est représenté par le titre du dictionnaire : s'il s'agit d'un dictionnaire de la langue générale ou d'un dictionnaire de pédagogie (d'apprentissage), on focalise sur la norme et l'on marque comme *savant* d'un côté, par exemple συστολική πίεση « tension systolique », ou comme *familier* de l'autre, ce qui dépasse les bords de la norme. Au contraire, s'il s'agit d'un dictionnaire de terminologie médicale, συστολική πίεση ne devrait porter aucune marque diaphasique (v. note 3).

Par conséquent, de notre point de vue, le problème du marquage ne se pose de façon accrue que pour les dictionnaires de la langue générale et les dictionnaires de pédagogie¹⁹. Mais comment pourrait-on améliorer la description linguistique au niveau des marques d'usage ? Comment décrire le vocabulaire de telle façon que tout locuteur s'y retrouve ? Est-il possible de doter de la marque adéquate chaque unité qui n'appartient pas à la norme ? Nous considérons que ce n'est pas possible de façon absolue, parce que la société elle-même n'est pas linguistiquement homogène et que les interactions langagières peuvent déclencher un changement de niveau de langue. Les descriptions présentes dans les dictionnaires sont le résultat d'un compromis, ce qui explique en partie la divergence dans l'estimation des lexicographes, qui reste largement intuitive. Par exemple, rien qu'en prenant en compte la variable de l'âge, telle unité linguistique peut paraître vulgaire pour des personnes âgées, tandis que pour des jeunes il s'agit d'une unité appartenant à la langue familière. Cette unité linguistique, de quelle marque doit-elle alors être dotée, vulgaire ou familière, dans un dictionnaire de la langue générale ? Ce qui revient à poser le problème d'une autre façon : la norme choisie, de quelle tranche d'âge reflète-t-elle la langue ? Y a-t-il une relation avec l'âge des lexicographes ?

4. PROPOSITION

Étant donné la non-homogénéité sociale, nous considérons que le marquage dans un dictionnaire de la langue générale ne peut être qu'un compromis du moins pour un certain nombre d'unités lexicales. Pour résoudre cette difficulté, nous faisons trois propositions :

17 Ici le terme de *dictionnaire de langue* ne renvoie pas au système linguistique mais s'oppose au *dictionnaire encyclopédique*.

18 V. aussi Rey & Delesalle (1979 : 6), qui reconnaissent qu'en théorie il y a deux réalités, la description du système et celle du système en fonctionnement, mais qu'en pratique tout dictionnaire procède à l'analyse des deux, en mettant diversement l'accent sur l'un ou sur l'autre.

19 Ce problème ne se pose pas pour les dictionnaires spécialisés, qui présentent la terminologie d'un domaine très ciblé.

1. S'orienter de la théorie à la pratique, c.-à-d. procéder à une définition univoque des marques sur la base de critères de délimitation, de façon à créer un dispositif de marquage explicite, affiné et cohérent, dont on retrouverait la trace dans les préfaces et les tables d'abréviations des dictionnaires (v. aussi Corbin 1989 : 678)²⁰.
2. Bannir des dictionnaires les jugements normatifs et ne conserver que les marques d'usage descriptives, sans aucune fonction normative ou puriste.
3. Procéder à une description plus détaillée des unités lexicales au niveau du marquage en dotant chacune de la marque correspondant à l'usage fait par les locuteurs de chacune des trois tranches d'âge auxquelles nous diviserions la société (jusqu'à 25 ans - jusqu'à 60 ans - au delà).

Cette dernière solution présente le désavantage d'augmenter la taille d'un dictionnaire papier, mais il est bien connu qu'elle est sans effet négatif pour un dictionnaire électronique. En plus, le support informatique ouvre des voies nouvelles pour le traitement de la marque, car les liens hypertextuels peuvent fournir :

- les différences diastématiques ou diaphasiques entre des unités synonymes ;
 - les différences diatechniques entre des unités appartenant, d'une part, au vocabulaire général et, de l'autre, à un vocabulaire spécialisé ;
 - des informations numériques sur la fréquence d'emploi d'une unité lexicale dans un corpus donné ainsi que sur son taux de différence par rapport à sa variante, p. ex. χούφτα – φούχτα « paume », ou entre la forme empruntée et son calque, p. ex. ασανσέρ – ανεγκυστήρας « ascenseur » ;
 - les différences diatopiques entre unités appartenant, d'une part, à la norme et, d'autre part, à différents dialectes – on peut même avoir recours au support visuel (cartes géographiques et images d'objets) et auditif pour mieux marquer les différences ;
 - les différences diachroniques entre des unités référant à des concepts voisins.
- Le support visuel peut aider, ici aussi, à noter :
- les différences entre des unités synonymes ayant trait au médium d'expression (oral ou écrit) ;
 - les différences entre unités se rapportant au type de discours (genre) ;
 - les différences connotatives entre unités (mélioratif ou péjoratif, intensif²¹).

CONCLUSION

Dans son œuvre maîtresse sur la lexicographie, Zgusta (1971 : 173) dit de la variation linguistique en synchronie qu'indiquer les spécificités lexicales de toutes ces variantes, bien que ça ne soit pas du tout chose facile, constitue une des tâches fondamentales du lexicographe. De même, J. Rey-Debove signale (1982 : xj) que la notation des niveaux de langue est délicate et peut prêter à des discussions mais qu'elle est, tout de même, nécessaire pour aider les usagers, surtout étrangers, à bien utiliser les mots.

20 Cette proposition n'est pas si facile à appliquer, car il faudrait d'abord délimiter les domaines de la sémantique et de la pragmatique.

21 Cela correspondrait à l'intensificateur *Magn* dans la théorie Sens-Texte de Mel'cuk.

L'existence des marques d'usage dans un dictionnaire rend explicite le phénomène de la variation linguistique, de la non-homogénéité d'une langue naturelle. C'est la raison pour laquelle nous jugeons hautement pertinente la présence des informations sur les marques dans un dictionnaire de la langue générale et un dictionnaire pédagogique.

Cependant, dans le discours lexicographique, il règne une incohérence et une incertitude non seulement sur les catégories et les critères utilisés par les lexicographes mais encore sur la formulation, le rôle et la place des marques. Les marques ne sont pas définies avec une rigueur suffisante pour constituer un métalangage opératoire, mais elles sont le produit d'une estimation intuitive de la part des lexicographes, ce qui nuit à la qualité des dictionnaires. Il faut donc une définition rigoureuse pour chaque marque d'usage ainsi qu'une cohérence dans leur distribution dans le lexique.

À l'issue de cette réflexion, nous espérons avoir contribué, d'une part, à mettre en lumière le manque de rigueur dans le système des marques d'usage utilisé dans les deux dictionnaires grecs et, de l'autre, à améliorer ce paysage confus par des propositions concrètes.

Les lexicographes grecs se sont efforcés de décrire l'usage effectif observé en utilisant les marques d'usage. Mais ils avaient à surmonter plusieurs types de difficultés, comme la situation linguistique de la Grèce après la chute de la junte militaire (1974) qui a entraîné l'abolition de la langue pure (*katharevoussa*) comme langue officielle de l'État²², le retard dans le domaine de la lexicographie grecque et dans la description du grec fondée sur les principes de la linguistique synchronique, etc.

Les différences entre les deux dictionnaires prouvent le caractère en partie arbitraire de ces décisions. Certes, l'amélioration objective et l'homogénéisation des descriptions dans les dictionnaires sont possibles et souhaitables (Rey & Delesalle 1979 : 26). Seul un travail sérieux préalable de la part des linguistes, des sociolinguistes, des métalxicographes et des lexicographes pourrait aboutir à un système clair et cohérent des marques d'usage, au moyen desquelles tout dictionnaire, papier et surtout électronique, deviendra un instrument de travail adapté aux besoins des usagers natifs ou étrangers.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ANASTASSIADIS-SYMÉONIDIS Anna, 1999, « Le métalangage lexicographique du grec moderne », in *Actes du colloque sur le 20^e anniversaire de l'institutionnalisation du grec moderne*, Athènes, Université d'Athènes, p. 117-135.
- ANASTASSIADIS-SYMÉONIDIS Anna et FLIATOURAS Asimakis, 2005, « La distinction [savant] et [non savant] en grec moderne : définition et classification », in *Actes du 6^e Congrès international de linguistique grecque*, vol. 1, Rethymon, Université de Crète, p. 110-120.
- ANASTASSIADIS-SYMÉONIDIS Anna et TZIVANOPOULOU Anastassia, 1984, « Le dictionnaire du grec moderne de l'Institut d'Études Néohelléniques (Fondation Manolis Triantaphyllidis) de l'Université Aristote de Thessaloniki », in

22 Avant 1974 et surtout entre 1967 et 1974, il était interdit de rédiger des dictionnaires de langue parlée. D'ailleurs, jusqu'à cette date, le métalangage lexicographique était en *katharevoussa*. Ce n'est qu'après 1974 que les premiers dictionnaires de langue parlée ont pu être rédigés, nécessitant la création d'une métalangue en langue parlée.

- A. Farmakides *et al.* (eds.), *The Teaching of Modern Greek in the English Speaking World*, Brookline [MA], Hellenic College Press, p. 93-100 (en grec).
- BABINIOTIS Giorgos, 1998, *Dictionnaire du grec moderne*, Athènes, Centre de Lexicologie.
- BURKHANOV Igor, 2003, "Pragmatic specifications : Usage indications, labels, exemples; dictionaries of style, dictionaries of collocations", in P. van Sterkenburg (ed.), *A Practical Guide to Lexicography*, Amsterdam, Benjamins, p. 102-113.
- COLLINOT André et MAZIÈRE Francine, 1987, « Un prêt-à-parler : le *Dictionnaire Universel* d'Antoine Furetière et sa postérité immédiate, le *Trévoux*. Une lecture du culturel dans le discours lexicographique », *Lexicographica*, 3, p. 51-75.
- CORBIN Pierre, 1989, « Les marques stylistiques / diastématiques dans le dictionnaire monolingue », dans F. J. Hausmann, O. Reichmann, H. E. Wiegand und L. Zgusta (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter, p. 673-680.
- CORBIN Pierre et CORBIN Danielle, 1980, « Le monde étrange des dictionnaires (1). Les "marques d'usage" dans le *Micro-Robert* », *Bulletin du Centre d'analyse du discours* 4 (« Synonymies »), Villeneuve d'Ascq, Presses Universitaires de Lille, p. 237-324.
- COSERIU Eugenio, 1976, « L'étude fonctionnelle du vocabulaire – Précis de lexicématique », *Cahiers de Lexicologie* 29 (2), p. 5-23.
- Dictionnaire de Proïas*, 1933, Athènes.
- Dictionnaire du grec standard*, 1998, Institut d'Études Néohelléniques, Thessalonique.
- DUBOIS Jean et DUBOIS Claude, 1971, *Introduction à la lexicographie – Le dictionnaire*, Paris, Larousse.
- GIRARDIN Chantal, 1979, « Contenu, usage social et interdits dans le dictionnaire », *Langue Française*, 43, p. 84-99.
- GIRARDIN Chantal, 1987, « Système des marques et connotations sociales dans quelques dictionnaires culturels français », *Lexicographica*, 3, p. 76-102.
- GLATIGNY Michel (éd.), 1990, *Lexique*, 9, *Les Marques d'usage dans les dictionnaires (XVII^e-XVIII^e siècles)*.
- GLATIGNY Michel, 1990, « Présentation : l'importance des marques d'usage », *Lexique*, 9, p. 7-16.
- HAENSCH Günther, WOLF Lothar, ETTINGER Stefan y WERNER Reinhold, 1982, *La Lexicografía*, Madrid, Editorial Gredos.
- HAUSMANN Franz Joseph, 1989, „Die Markierung im allgemeinen einsprachigen Wörterbuch: eine Übersicht“, in F. J. Hausmann, O. Reichmann, H. E. Wiegand und L. Zgusta (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter, p. 649-657.

- IMBS Paul (éd.), 1971, *Trésor de la langue française*, tome I, Paris, Éditions du CNRS.
- LANDAU Sidney, 1989, *Dictionaries. The Art and Craft of Lexicography*, Cambridge, Cambridge University Press.
- MAZIÈRE Francine, 1990, « Les marques de *fabrique*. Marquage et marques de domaine dans les dictionnaires classiques, du *Furetière* aux *Trévoux* », *Lexique*, 9, p. 89-111.
- MEL'ČUK Igor, 1993, *Cours de morphologie générale (théorique et descriptive)*, Montréal, Les Presses de l'Université de Montréal, et Paris, CNRS Éditions.
- NORRI Juhani, 1996, "Regional labels in some British and American dictionaries", *International Journal of Lexicography*, 9 (1), p. 1-29.
- NORRI Juhani, 2000, "Labelling of derogatory words in some British and American dictionaries", *International Journal of Lexicography*, 13 (2), p. 71-106.
- PETIT Gérard, 2010, *La Dénomination : approches linguistique et terminologique* (mémoire d'HDR de l'Université Paris 13), Louvain-la-Neuve, Peeters (collection « Bibliothèque de l'Information grammaticale »).
- REY Alain (éd.), 1967, 1969, *Le Petit Robert*, Paris, Société du Nouveau Littré.
- REY-DEBOVE Josette, 1971, *Étude linguistique et sémiotique des dictionnaires français contemporains*, La Haye, Mouton.
- REY-DEBOVE Josette (éd.), 1982, *Le Robert Méthodique - Dictionnaire méthodique du français actuel*, Paris, Le Robert.
- REY Alain et DELESALLE Simone, 1979, *Langue Française*, 43, Paris, Larousse, p. 4-26.
- VERKUYL Henk, JANSSEN Maarten and JANSEN Frank, 2003, "The codification of usage by labels", in P. van Sterkenburg (ed.), *A Practical Guide to Lexicography*, Amsterdam, Benjamins, p. 297-311.
- ZGUSTA Ladislav, 1971, *Manual of Lexicography*, The Hague, Mouton.

PREMIÈRE PARTIE
ASPECT SÉMANTIQUE

II.

ÉQUIVALENCE

CLASSIFICATION DES FORMES SENTENCIEUSES ET TRADUCTION D'ÉNONCÉS PARÉMIQUES

Jean-Claude ANSCOMBRE
CNRS

Résumé — Cette étude expose certains problèmes auxquels on se trouve confronté dès lors qu'on se frotte à la traduction des formes sentencieuses, et plus particulièrement des proverbes. Étant donné que nous procédons à l'élaboration d'un dictionnaire bilingue espagnol-français de formes sentencieuses, nous sommes donc confronté *a diario* au problème de la recherche d'équivalences d'énoncés parémiques entre le français et l'espagnol. Cet article est le fruit des réflexions et des conclusions théoriques et pratiques auxquelles nous avons abouti. Le problème est celui bien connu de tous les traducteurs — qu'ils soient professionnels ou comme nous de simples amateurs: la traduction de tout ce qui a un aspect formulaire, renvoie non seulement à la langue, mais également à un savoir commun qu'il s'agit donc de restituer au mieux lors du passage de la langue-source à la langue-cible. Et les ouvrages bilingues sur de tels sujets sont la plupart du temps inutilisables, les équivalences proposées étant généralement grossières, au mieux médiocres. Phénomène qui est dû à des principes traductologiques implicites bien que toujours présents, en particulier : (a) Ce qui a été autrefois un proverbe est toujours un proverbe ; (b) Étant donné un proverbe d'une langue L_1 donnée, tout énoncé d'une langue L_2 qui ressemble et qui est un proverbe fournit une équivalence valable du premier ; (c) C'est dans les compilations et dictionnaires antérieurs que l'on trouve des proverbes « qui se ressemblent ». Après avoir présenté un certain nombre d'exemples de traductions que nous considérons comme « ratées », nous montrerons l'origine de l'erreur commise. Dans un second temps, nous essaierons d'introduire quelques principes méthodologiques relatifs à l'étude des proverbes et formes parémiques, et montrerons comment les utiliser pour éviter les erreurs ou, à tout le moins, améliorer l'équivalence. Le problème de fond peut être résumé de la façon suivante : Quelles marques portées par une forme sentencieuse doivent impérativement être considérées pour assurer une bonne équivalence ?

Mots clés — proverbes, traduction culturelle, dictionnaire espagnol / français, méthodologie, traductologie

Abstract — This study discusses various problems that translators face when dealing with proverbs, looking at the specific case of writing a bilingual Spanish

/ French proverb dictionary. Indeed translating any phrase or proverb involves not only an excellent grasp of a language but also a shared knowledge. Owing to faulty methodology, most bilingual proverb dictionaries give mediocre equivalencies. For example, there seems to be, the belief that what was once a proverb is always a proverb, or that any sentence in L2 similar to L1 will give an acceptable equivalent. Working from several “bad” translations, the author explains which methodological principles should be applied when translating phrases and proverbs and how they should be applied in order to improve this field in translation studies.

Keywords — translation of proverbs, cultural transfer, Spanish-French proverb dictionary, methodological principles, translation theory

INTRODUCTION

Nous voudrions dans cette brève étude exposer certains problèmes auxquels on se trouve confronté dès lors qu'on se frotte à la traduction de ce type particulier de construction que sont les formes sentencieuses, et plus particulièrement les proverbes. Cette préoccupation est actuellement la nôtre puisque nous procédons à l'élaboration d'un dictionnaire bilingue espagnol-français de formes sentencieuses. Nous sommes donc confronté *a diario* au problème de la recherche d'équivalences d'énoncés parémiques entre le français et l'espagnol, et c'est le fruit des réflexions et des conclusions théoriques et pratiques auxquelles nous avons abouti que nous livrons ici.

Le problème est celui bien connu de tous les traducteurs – qu'ils soient professionnels ou comme nous de simples amateurs : la traduction de tout ce qui a un aspect formulaire renvoie non seulement à la langue, mais également à un savoir commun¹ qu'il s'agit donc de restituer au mieux lors du passage de la langue-source à la langue-cible. Et les ouvrages bilingues sur de tels sujets sont la plupart du temps inutilisables, les équivalences proposées étant généralement grossières, au mieux médiocres. Phénomène qui est dû à des principes traductologiques implicites bien que toujours présents, en particulier : (a) ce qui a été autrefois un proverbe est toujours un proverbe ; (b) étant donné un proverbe d'une langue L_1 donnée, tout énoncé d'une langue L_2 qui ressemble et qui est un proverbe fournit une équivalence valable du premier ; (c) c'est dans les compilations et dictionnaires antérieurs que l'on trouve des proverbes « qui se ressemblent ».

Nous commencerons par un certain nombre d'exemples de traductions selon moi « ratées », et montrerons l'origine de l'erreur commise. Dans un second temps, nous essaierons d'introduire quelques principes méthodologiques relatifs à l'étude des proverbes et formes parémiques, et montrerons comment les utiliser pour éviter les erreurs ou, à tout le moins, améliorer l'équivalence. Le problème de fond peut être résumé de la façon suivante : Quelles marques portées par une forme sentencieuse doivent impérativement être considérées pour assurer une bonne équivalence ?

¹ Le problème du savoir commun est qu'il s'agit très souvent d'un non-dit auquel il s'agit de trouver un non-dit correspondant dans la langue cible. Le problème est particulièrement aigu lorsqu'on est en présence de deux cultures très dissemblables.

1. LA TRADUCTION DE PROVERBES : QUELQUES CAS PROBLÉMATIQUES

Nous tirerons le premier cas du recueil *1001 Refranes españolas y su correspondencia en alemán, árabe, francés, inglés, italiano, polaco, provenzal y ruso*. Ce recueil traduit le proverbe espagnol *Agua pasada no mueve molino*, litt. « Eau passée ne meut pas le moulin », par « Le moulin ne meut pas avec l'eau passée en bas ». Cette bizarre traduction frappe d'entrée par son caractère incorrect en français contemporain : on devrait dire « Le moulin ne se meut pas sous l'action de l'eau qui se trouve en aval », entre autres possibilités. De plus, cet énoncé ne fait pas partie du corpus des proverbes français contemporains, ni même anciens. Erreur qui ne fait en réalité que reproduire une erreur antérieure, présente dans Maloux, à savoir *Le moulin ne moud pas avec l'eau passée en bas*, proverbe en fait repris de *Adagios and Sententious Proverbs*, de T. Draxe, 1616. La forme anglaise actuelle est *The mill cannot grind with the water that is past*. On trouve là une belle illustration de (c) : utilisation non contrôlée de dictionnaires et recueils antérieurs.

Le second exemple est celui de l'équivalence entre *Piedra movediza, nunca moho la cobija* et *Pierre qui roule n'amasse pas mousse*, dont les traductions littérales sont très proches, et qui sont présentées partout comme équivalentes. Or malheureusement, cette phrase sentencieuse a un sens toujours négatif en français, et généralement positif en espagnol, sauf dans certaines régions. La version anglaise *A rolling stone gathers no moss* pouvant avoir l'un ou l'autre sens, selon les régions, comme déjà signalé par Milner (1969). On a là un exemple typique de la thèse (b) : l'équivalence par ressemblance. Notons que l'équivalence par ressemblance fait des formes sentencieuses de simples entrées lexicales, en oubliant que leur rôle essentiel est l'articulation du discours. On ne peut donc faire table rase des contextes d'apparition des formes sentencieuses, puisqu'ils sont l'indice du sens de ces formes.

Troisième exemple : *En boca cerrada no entran moscas*, que N. Urdíroz Villanueva² traduit par *En bouche close n'entre mouche*, traduction dont elle estime qu'il s'agit là d'un homologue tant pour la forme que pour le sens. Selon elle, la version originale, très proche, apparaît dans Carmen, et Mérimée la fait remonter à un proverbe gitan. En fait, Mérimée propose *En bouche close n'entre point mouche* comme la traduction d'un proverbe gitan, ce qui laisse ouvert le problème de savoir si, à l'époque, il s'agissait déjà d'un proverbe. On peut cependant se demander si c'est le cas ou si Mérimée n'a pas tout simplement traduit le proverbe espagnol – n'oublions pas qu'il fut le précepteur d'Eugénie de Montijo. À époque contemporaine, cette forme n'existe plus et ferait sourire si elle était utilisée. Par ailleurs, est attestée en 1610 dans le *Florilegium* de Gruter une forme *En bouche close n'entre mouche*, qui n'est mentionnée nulle part par la suite, ce qui amène à se poser la question d'un possible hapax, et illustre de toutes façons le point (a).

Quatrième et dernier exemple, celui de la traduction (d'ailleurs fort bonne) d'un roman de science fiction³, dans lequel on découvre la phrase suivante :

2 «La presencia de dichos y refranes en *La familia de Pascual Duarte*», *Paremia*, 8, 1999 : 517-520.

3 Philip José FARMER, *Des Rapports étranges*, éd. J'ai lu, n° 712, trad. de Michel Deutsch, p. 98.

« ... Il m'a flanqué la plus belle râclée de ma vie... paix à son âme, ajouta-t-il...
Sans fouet, enfant gâté était sa maxime favorite... »

Il s'agit visiblement d'une adaptation du proverbe anglais *Spare the rod, spoil the child*, pour lequel le traducteur n'a pas su trouver d'équivalent. Il en a donc forgé un, nous verrons plus loin comment, ainsi qu'une méthode pour trouver un équivalent authentique.

2. FORMES SENTENCIEUSES, PROVERBES, ETC. : QUELQUES DÉFINITIONS

La plupart des auteurs se consacrant à l'étude des formes sentencieuses (proverbes, dictons, adages, sentences, etc.) ne se posent guère de questions « théoriques », et le domaine se caractérise par une terminologie aussi floue qu'abondante. Parmi les termes le plus fréquemment rencontrés, je ne mentionnerai ici que les principaux, à savoir : *adage*, *apophtegme*, *aphorisme*, *dicton*, *maxime*, *précepte*, *sentence proverbe*, etc.

Les recueils contemporains, en total désaccord entre eux, nous proposent la plupart du temps des classements fantaisistes sous des appellations encore plus fantaisistes. Considérons en effet la très célèbre forme sentencieuse *Une hirondelle ne fait pas le printemps*. DesRuisseaux la classe dans les proverbes (comme le *Grand Robert* et le *TLF*), alors que Pierron le considère comme un dicton. Quant à Maloux, il le classe dans les proverbes du grec ancien – bien que cette forme sentencieuse existe dans à peu près toutes les langues indo-européennes contemporaines, comme on le verra. Delacourt y voit une maxime, Djavadi parle de dicton météorologique, et prudents, Rey & Chantreau ne le classent pas, non plus que Lis & Barbier, qui y voient une « observation ». Montreynaud, Pierron & Suzzoni renvoient pour l'origine à Le Roux de Lincy, qui renvoie à son tour à un auteur du XVI^e s., et ce malgré l'existence du modèle latin pourtant bien connu *Una hirundo non effecit ver*. Certains recueils présentent des traductions de proverbes étrangers, selon des choix discutables et des traductions souvent encore plus discutables, mais c'est bien avec l'hypothèse implicite que la traduction d'un proverbe d'une langue est un proverbe dans l'autre langue, par la force des choses. Hypothèse qui n'est qu'un cas particulier de l'hypothèse plus générale qu'un proverbe, quelles que soient sa région et son époque d'origine, sera à jamais proverbe : *Once a proverb, always a proverb*. D'où des équivalences douteuses, ainsi celle-ci, trouvée dans un recueil récent ((Montreynaud, Pierron & Suzzoni 1980: 374) : *L'ormeau ne peut donner des poires*, traduction quelque peu fantaisiste du proverbe espagnol *No hay que pedir peras al olmo*, qui serait rendu de façon à mon avis plus satisfaisante par *Il ne faut pas demander l'impossible*, voire *À l'impossible nul n'est tenu*.

La critique est facile, dit-on. Ce désordre est l'indice d'un problème de fond, auquel il convient de remédier. L'idée de départ sera que, dans la mesure du possible, le traducteur doit s'assurer qu'à une forme sentencieuse de la langue-source il fera correspondre une forme sentencieuse de la langue-cible, avec une correspondance sémantico-pragmatique aussi fidèle que possible – exigence minimale de toute traduction. Il nous faut donc fournir des critères d'identification des formes sentencieuses, voire même de sous-classes de telles formes, l'idée étant que les propriétés des formes sentencieuses se manifestent par des

marques de surface qui sont autant d'indications sur la nature sentencieuse exacte de la forme considérée.

Une première caractérisation sera que les phénomènes qui nous occupent ici ont trait à des constructions phrastiques, qui ont une propriété les distinguant d'autres constructions phrastiques : à l'intérieur d'un discours donné, elles sont généralement mobiles, et font donc partie des *phrases autonomes*. On peut par exemple, dans *Les dés sont jetés, inutile d'insister*, rejeter la phrase sentencieuse en position finale sans changement notable de sens : *Inutile d'insister, les dés sont jetés*. Seconde remarque : toutes les phrases autonomes ne méritent pas le statut de phrase sentencieuse : ainsi la phrase *Les voitures ont quatre roues* est également autonome⁴, et ne semble guère sentencieuse. À l'intérieur donc des phrases autonomes, nous distinguerons les *phrases sentencieuses* par le critère de combinabilité avec la tournure *Comme dit X*, *X* étant l'auteur présumé de la phrase autonome. Par exemple :

- (1) *Comme (le) dit La Rochefoucauld, on ne blâme le vice et on ne loue la vertu que par intérêt.*
- (2) *Comme on dit, au royaume des aveugles, les borgnes sont rois.*

Ce critère distingue bien deux sous-classes parmi les phrases autonomes, puisqu'on n'aura jamais :

- (3) **Comme on dit, les voitures ont quatre roues.*

Bien qu'également générique, la phrase *Les voitures ont quatre roues* n'est pas une phrase sentencieuse selon ce critère. Une telle remarque n'est pas sans intérêt : elle attire en effet l'attention sur le fait qu'une phrase d'un texte-source donné est sentencieuse car accompagnée d'un commentaire méta-linguistique qui le prouve. Et qu'on doit par conséquent lui fournir une équivalence sentencieuse dans la langue-cible. Or de tels commentaires « méta-linguistiques »⁵ ne sont pas rares dans les textes, ainsi :

- (4) «...El hombre tortura y mata porque es lo suyo. Le gusta. —¿Lobo para el hombre, como dicen los filósofos?...» (A. Pérez-Reverte, *El pintor de batallas*, Madrid, Alfaguara, 2006, p. 106).
- (5) ... Bien évidemment, l'éditeur de Richmond a su rapidement réagir, en mettant un patch de sécurité – téléchargeable sur Internet – à la disposition des utilisateurs. Et *comme le voudrait le proverbe*, « Faute avouée est à moitié pardonnée... » (*Windows*, n° 96, février 2002, p. 13).
- (6) ... *As they say*, no fool like an old fool... (L. Barnea, *Reported Missing*, 1979, p. xvij).

Tout traducteur, même ignorant le sens exact de *Faute avouée est à moitié pardonnée*, saura d'entrée, en vertu de la présence de *comme le voudrait le proverbe*, qu'il s'agira d'en trouver une traduction sentencieuse.

Parmi les phrases sentencieuses ainsi définies – autonomes et combinables avec *comme dit X* –, nous distinguerons deux types. Le premier correspond à un

4 Il me faut quatre pneus, les voitures ont quatre roues / Les voitures ont quatre roues, il me faut quatre pneus.

5 Outre *comme on dit*, les plus courants sont *comme dit la sagesse populaire*, *comme dit l'autre*, *comme on dit chez moi*, *dit-on*, et *On sait bien que*. De façon un peu plus spécifique, on trouve également en abondance des tournures comme *comme le dit un vieux proverbe*.

auteur précis – ainsi La Rochefoucauld dans l'exemple ci-dessus : c'est le domaine des maximes, sentences, morales, qui puisent habituellement dans l'histoire et la littérature attachées à la langue considérée. Nous réserverons ici nos efforts au second cas, celui de l'auteur anonyme, que les linguistes⁶ ont baptisé du curieux nom de *ON*-locuteur. Ce *ON*-locuteur est fréquemment désigné par la *sagesse des nations*, la *sagesse populaire*, le *bon sens populaire*, etc. Voici quelques formes sentencieuses à auteur anonyme, empruntées à diverses langues :

- (7) (fr.) *La mariée était trop belle.*
- (8) (fr.) *Mieux vaut un mauvais accord qu'un bon procès.*
- (9) (angl.) *A friend in need is a friend indeed.*
- (10) (angl.) *It takes two to tango.*
- (11) (esp.) *No hay mal que por bien no venga.*
- (12) (esp.) *No hay moros en la costa.*
- (13) (all.) *Eine Schwalbe macht noch keinen Sommer.*
- (14) (all.) *Wie der Vater, so der Sohn.*
- (15) (cat.) *Tal farás, tal trobarás.*
- (16) (cat.) *Una flor no fa estiu.*

Elles se combinent aisément – on le vérifiera – avec les marqueurs du type de *Comme on dit* :

- (17) *Comme on dit, mieux vaut un mauvais accord qu'un bon procès.*
- (18) *As they say, it takes two to tango.*

À l'intérieur des formes sentencieuses à *ON*-locuteur, nous distinguerons encore deux sous-classes. Celle des formes sentencieuses qui sont génériques – elles représentent une vérité valable *extra tempore*, et celle des formes sentencieuses qui ne le sont pas. À ces dernières on réserve le nom de *phrases situationnelles* (Kleiber 1989). Elles qualifient directement une situation, et en font partie des phrases comme *La mariée est trop belle*, *Un ange passe*, *Il y a de l'eau dans le gaz*, *C'est bonnet blanc et blanc bonnet*, *Les dés sont jetés*, *Il a coulé de l'eau sous les ponts*, *Les carottes sont cuites*, etc. Elles sont en fait événementielles, et portent ou peuvent porter des marques circonstanciellles, ce qui n'est pas le cas des sentencieuses génériques. On comparera de ce point de vue (19) et (20) à (21) et (22) :

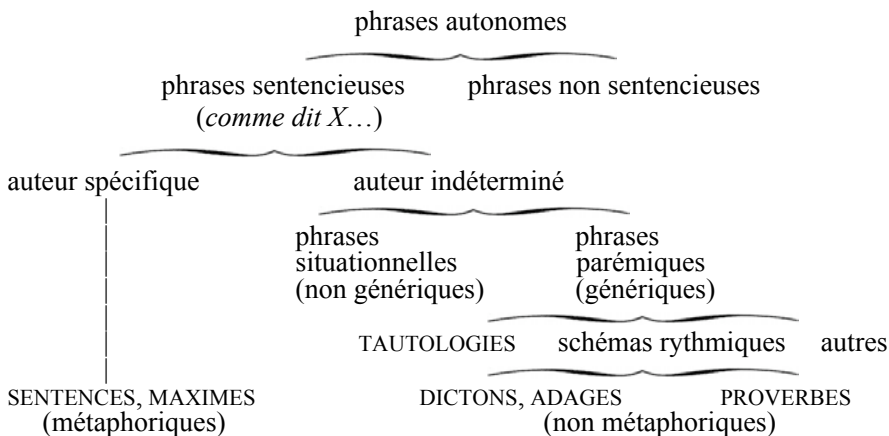
- (19) *Attention, dans cette maison les murs ont des oreilles.*
- (20) *Tout le monde s'est tu. Un ange est passé.*
- (21) **Attention, dans cette maison, une hirondelle ne fait pas le printemps.*
- (22) **Tout le monde s'est tu. Qui n'a dit mot a consenti.*

À l'ensemble des phrases sentencieuses ayant un *ON*-locuteur et de nature générique, nous réserverons le nom de *phrases parémiques*. Catégorie elle-même divisée en plusieurs sous-classes, à savoir :

- (a) La sous-classe des *tautologies*, phrases comme *Un sou est un sou*, angl. *Duty is duty*, all. *Mann ist Mann*, cat. *El diner és diner*, esp. *Un día es un día*, ital. *Domenica è sempre domenica*, etc. On les reconnaît essentiellement à une structure superficielle très souvent proche de « p est p ».

- (b) La sous-classe des phrases parémiques exhibant certains schémas rythmiques, elle-même divisée en deux sous-classes, selon que la phrase est métaphorique ou non. La première classe, celle des [proverbe], correspond *grosso modo* aux « proverbes » au sens habituel, la seconde est celle des [adage] et des [dicton]⁷.
- (c) Le reste, i.e. la sous-classe des phrases parémiques n'ayant pas d'autre(s) propriété(s) saillante(s). Ainsi *Les apparences sont trompeuses*, ou encore *Chacun voit midi à sa porte*.

Sans entrer dans des détails techniques, voici quelques exemples. Le bien connu *Un mauvais arrangement vaut mieux qu'un bon procès* se combine sans problème avec *Comme on dit*, et est assez notoirement générique – on peut le combiner avec *souvent* / *toujours* sans en trahir l'esprit. Il s'agit donc d'une phrase parémique, dont la structure superficielle n'est pas celle d'une tautologie. Le découpage *Un mauvais arrangement* / *Vaut mieux qu'un bon procès* (suggéré par le parallélisme *mauvais arrangement* / *bon procès* dans GN et GV) montre une structure rythmique 6 / 6. Il n'est pas métaphorique, il s'agit donc d'un [adage]. *La pluie du matin n'arrête pas le pèlerin* découpée selon GN / GV comme précédemment montre une structure 5 / 7⁸ et une rime -in / -in qui en fait cette fois un [dicton]. Soit maintenant *Qui va à la chasse / perd sa place* : il exhibe une rime -as / -as sur une structure 5 / 3⁹, et est métaphorique : c'est bien un [proverbe]. Enfin, *Chacun voit midi à sa porte* n'a pas de caractéristique particulière en dehors de la métaphoricité : c'est une simple phrase parémique. Notons enfin que les cloisons entre les différentes catégories ne sont pas étanches, et qu'il arrive qu'on passe de l'une à l'autre :



7 J'utiliserai *grosso modo* [adage] pour les phrases parémiques non métaphoriques à contenu moral ou juridique, [dicton] pour celles concernant le rapport avec la nature. La notation [C], que j'utilise depuis plusieurs années, représente le concept défini sur la base de critères linguistiques, et non le mot de la langue, qui ne peut en aucune façon être considéré comme un concept scientifique. Pour plus de détails, v. Anscombe 2006a.

8 Une telle structure – combinaison d'heptasyllabes et de pentasyllabes est très fréquente dans le domaine parémique, mais également dans le domaine poétique: ainsi le *haïku* et le *tanka* japonais, mais aussi la *seguidilla* espagnole. En français, le phénomène est moins net, peut-être à cause de l'extraordinaire succès de l'alexandrin dodécasyllabique à la Renaissance.

9 Les combinaisons de ce type sont très communes, les plus courantes étant 7/5 , 6/4, et 5/3.

Le problème qui nous occupe ici peut alors se formuler de la façon suivante : Étant donné une forme sentencieuse *S* appartenant à une sous-classe donnée d'une langue-source, comment lui faire correspondre au mieux une forme sentencieuse *S'* d'une langue-cible ? En d'autres termes, comment faire – et est-ce toujours possible – pour que la traduction proposée reste dans la même sous-classe sentencieuse ? Comme annoncé, nous utiliserons principalement l'espagnol et le français comme source et cible.

3. À LA RECHERCHE DE L'ÉQUIVALENCE PERDUE

3.1 LE FONDS COMMUN

Le premier réflexe de tout traducteur non professionnel, lorsqu'il lui faut fournir une équivalence d'une forme sentencieuse, est bien entendu de chercher « quelque chose qui ressemble ». Cette démarche est quelque peu naïve : il est assez exceptionnel que traduire se réduise à calquer, y compris et surtout dans le domaine sentencieux, et pour certaines raisons qui apparaîtront plus loin. Elle n'est cependant pas totalement dépourvue de sens, du moins à l'intérieur du champ indo-européen. En effet, l'utilisation massive du latin comme langue véhiculaire jusqu'au moyen-âge ainsi que l'importance de la culture classique gréco-latine dans notre enseignement ont fait que le domaine sentencieux indo-européen contemporain provient majoritairement d'un fonds commun latin, éventuellement issu du grec classique, que ce soit au niveau de l'idée de base, ou même parfois de la forme. Voici un cas assez étonnant, où les langues indo-européennes modernes semblent avoir pratiquement copié le modèle latin. Il s'agit du déjà cité *Una hirundo non effecit ver* « Une hirondelle ne fait pas le printemps », qui outre l'équivalent français a donné :

- (all.) *Eine Schwalbe macht noch keinen Sommer*,
- (angl.) *One swallow does not make a summer*,
- (cat.) *Una orenata no fa estiu*¹⁰,
- (esp.) *Una golondrina no hace verano*,
- (gall.) *Unha andorinha soa non fai verán*,
- (grec moderne) *Eva xelidóni den férνει την άνοιξη*¹¹,
- (ital.) *Una rondine non fa primavera*,
- (port.) *Uma andorinha não faz primavera*,
- (russe) *Odna lastochka vesnij ne delaet*,
- (suéd.) *En svala gör ingen sommar*.

Ce cas n'est nullement exceptionnel, il y en a d'autres : ainsi, sur le modèle latin de *Canes qui plurimum latrant, perraro mordent*, on a formé :

- (ang.) *A barking dog never bites*,

¹⁰ Le catalan lui préfère cependant *Una flor no fa estiu*.

¹¹ C'est vrai aussi pour le bulgare, le serbo-croate, l'ukrainien et même le finnois. Le basque semble avoir également suivi ce modèle: *Ainhara batek ez du udaberria egiten*, lit. « hirondelle / une / non / il a / printemps le / fait » (dialecte labourdin). Les variations selon les langues entre « printemps » et « été » ne viennent pas de l'observation de variations climatologiques, mais d'une erreur de traduction. En latin, *prima vere* était la première partie du printemps, et *veranum tempus* la seconde partie, l'été étant désigné par *aestivum tempus*. D'où le doublet espagnol *verano / estío*.

- (cat.) *Gos que lladra no mossega,*
- (esp.) *Perro ladrador poco mordedor,*
- (fr.) *Chien qui aboie ne mord pas,*
- (it.) *Can che abbaia non morde,*
- (port.) *Cão que ladra não morde,*
- etc.

En cas donc de « panne de traduction », on peut effectivement penser à aller chercher l'original latin correspondant à la forme sentencieuse de la langue source, lequel peut suggérer une équivalence dans la langue cible. Voici un exemple où effectivement ce procédé donne un résultat positif. Considérons la très connue forme sentencieuse espagnole *Zapatero, a tus zapatos*, litt. « Cordonnier, à tes chaussures », et supposons pour la démonstration que nous ayons un certain mal à trouver une correspondance en français langue-cible. L'original latin – si l'on en croit du moins Pline l'Ancien, est *Ne sutor ultra crepidam*, « Cordonnier, (ne va) pas au-delà de la sandale », version latine de la réplique (grecque) supposée du peintre Apeles à un cordonnier qui critiquait un de ses tableaux. Le sens devient alors plus clair : le cordonnier n'y joue qu'un rôle accessoire, l'ensemble étant une version polie de « Occupe-toi de tes affaires », ou encore de « Femme, à tes casseroles ». Il s'agit d'une forme sentencieuse destinée à faire savoir que la compétence de chacun ne peut ni ne doit dépasser sa sphère d'activité. Or la sphère d'activité comporte essentiellement le métier, ce qui suggère immédiatement la correspondance *À chacun son métier, et les vaches seront bien gardées*.

Ce procédé, s'il n'est pas à rejeter, possède cependant ses limites. La ressemblance de surface peut n'être précisément que de surface – les apparences sont parfois trompeuses – même si les deux formes sentencieuses comparées proviennent d'un unique modèle latin. Reprenons ainsi l'exemple ci-dessus *Chien qui aboie ne mord pas*, l'origine latine en étant *Canes qui plurimum latrant, perraro mordent*. On aurait donc tendance à voir dans l'espagnol *Perro ladrador, poco mordedor* la parfaite équivalence cherchée. Et on aurait tort : la forme sentencieuse espagnole possède en effet deux sens, d'une part « une menace bruyante est rarement une véritable menace », sens attesté entre autres dans les zones nord et est de l'Espagne. Mais aussi « ceux qui parlent beaucoup agissent rarement », sens commun dans la zone de la capitale¹², et qui correspond en fait au français *Les conseillers ne sont pas les payeurs*, et surtout à l'anglais *All talk and no action*.

On voit où est le problème : certes, c'est bien d'une certaine façon une ressemblance que cherche le traducteur. Mais nous n'avons pas défini ce qu'était une *ressemblance*, ou du moins, nous avons limité son champ d'application à la seule structure superficielle des phrases sentencieuses concernées.

Nous allons donc nous efforcer dorénavant de délimiter la notion de ressemblance / équivalence, aussi indispensable qu'elle est floue. Nous verrons qu'on peut en fait définir différents types d'équivalence possibles, dont nous illustrerons l'application mais aussi les limitations.

¹² C'est d'ailleurs le sens officiel que lui donne la Real Academia Española dans son *Diccionario*.

3.2 L'ÉQUIVALENCE CATÉGORIELLE

Une première idée de l'équivalence, et qui vient assez naturellement à l'esprit, est l'idée de l'équivalence catégorielle. Elle signifie qu'à une forme sentencieuse d'une certaine catégorie, on doit s'efforcer de faire correspondre une forme sentencieuse de la même catégorie, et que ce faisant, on sert mieux la cause de la traduction. Outre l'idée de ressemblance, c'est aussi l'idée de fonctionnalité semblable qui prévaut ici. En effet, le fonctionnement des différentes catégories sentencieuses n'est pas le même. Ainsi, les phrases situationnelles comme *Il a passé / coulé de l'eau sous les ponts* ou *C'est bonnet blanc et blanc bonnet* se contentent de la simple appréciation d'une situation, alors que les tautologies ont souvent un côté prescriptif ou normatif, ainsi *Quand faut y aller, faut y aller*, *Une promesse est une promesse*, *Donner, c'est donner*, etc. Prescriptifs aussi apparemment, des adages comme *L'exactitude est la politesse des rois*, *Un mauvais arrangement vaut mieux qu'un bon procès* ou *Jeux de mains, jeux de vilains*, ainsi que les dictons agricoles *À la Saint Rémi, cueille tes fruits*, ou *À la Sainte Catherine, tout bois prend racine*. En revanche, les dictons météorologiques se présentent comme des observations : *Petite pluie abat grand vent*, ou *Noël au balcon*, *Pâques aux tisons*, *À la Sainte Luce, le jour croît du saut d'une puce*, sans dimension prescriptive. Les proverbes enfin présentent ce qu'il est convenu d'appeler des *doxa*, i.e. des schémas de raisonnement¹³. Concluons : pour ne pas se tromper de fonctionnement lors de la recherche d'équivalences, mieux vaut les chercher à l'intérieur de la même catégorie. Il ne manque d'ailleurs pas, dans le domaine indo-européen qui m'occupe ici, de cas où l'équivalence peut effectivement avoir lieu à l'intérieur d'une même catégorie. Soit par exemple la phrase sentencieuse *Il a coulé de l'eau sous les ponts*. Il s'agit d'une phrase situationnelle, reconnaissable au fait qu'elle peut être combinée avec un déictique comme *Là*, et qu'il y a possibilité de variations temporelles :

- (23) ... *D'ici* à ce que l'ensemble des ménages français envisage systématiquement d'installer un PC complet dans le salon, *de l'eau va encore couler sous les ponts*... (*Windows*, mars 2005, n° 130, p. 49).

Or l'espagnol possède une phrase sentencieuse de même sens, à savoir *Ha llovido desde entonces*, qui est également situationnelle, comme le montre l'attestation suivante :

- (24) ...Aristóteles, prosiguió imperturbable, nunca se limitó a exponer lo que sucedía, sino que buscó el porqué... Lo que pasa es que *desde entonces, ha llovido mucho*... (A. Pérez-Reverte, *El pintor de batallas*, Madrid, Alfaguara, 2006, p. 120)¹⁴

où l'expression en question est accompagnée (comme c'est fréquemment le cas) du déictique *entonces* « alors », et de plus fléchissable à l'imparfait ou au futur. Malheureusement, tout n'est pas pour le mieux dans le meilleur des mondes, et

13 Désireux de ne pas entrer dans des détails de technique linguistique qui n'ont pas leur place ici, je me contenterai d'un survol rapide des différents modes de fonctionnement. Les faits sont plus compliqués, et on peut opérer des distinctions plus fines.

14 « Aristote, poursuivit-il sans se troubler, ne s'est jamais borné à expliquer ce qui se passait, mais a au contraire cherché le pourquoi des choses. Ce qu'il y a, c'est que depuis lors il a coulé de l'eau sous les ponts. »

ce cas idyllique ne saurait être considéré que comme une heureuse exception. Bien souvent, le traducteur de phrases sentencieuses devra se résoudre à changer de catégorie s'il veut fournir une équivalence plausible. Considérons ainsi la phrase sentencieuse espagnole mentionnée plus haut et fort courante, *Agua pasada no mueve molino*, qui signifie *grosso modo* que les choses passées ont perdu toute pertinence. Il s'agit d'une expression métaphorique, et son schéma rythmique /'-/-'--/-'--/-'--, presque complètement régulier, montre qu'il s'agit d'un [proverbe]. Le français ne possède rien d'équivalent, et le traducteur devra donc choisir entre deux possibilités : traduire notre phrase espagnole par *Le passé est le passé*, qui conserve l'idée, mais perd la métaphore de l'espagnol, et de plus change de catégorie : la phrase française est en effet une [tautologie] qui, de plus, existe aussi en espagnol : *El pasado, pasado está*. Ou alors se résigner à rendre l'original par *Laissons les morts enterrer les morts*, qui est métaphorique et présente une pseudo-rime (*morts / morts*), mais n'a pas la généralité de la tournure espagnole, qui peut s'appliquer à des circonstances plus frivoles que sa correspondance française. Un autre problème réside en ce que bien souvent dans le domaine sentencieux, l'équivalence y compris catégorielle peut recéler des pièges redoutables. Soit par exemple le cas de *Un sou est un sou*, visiblement une [tautologie]. À première vue, rien de plus facile : l'espagnol possède justement une phrase sentencieuse, *La pela es la pela*¹⁵, qui semble convenir parfaitement à tous points de vue. C'est une [tautologie], apparemment de même sens que l'original, et qui appartient au registre familier, tout comme l'original. En fait, la forme sentencieuse espagnole se prononce habituellement avec un *l* rétroflexe à l'imitation de la prononciation catalane, et pour se moquer des catalans et de leur amour supposé de l'argent. Une traduction plus exacte de *La pela es la pela* serait donc en fait *Un chou c'est un chou*¹⁶, ce qui nous laisse sans véritable équivalent pour *Un sou est un sou*. On peut cependant s'en tirer avec *Il n'y a pas de petit profit*, qui présente le défaut de ne pas être une [tautologie], mais une simple phrase parémique sans caractéristiques particulières.

On voit ainsi apparaître le fond du problème : certes, l'équivalence catégorielle est importante, mais elle est, à y regarder de plus près, d'une plus grande complexité que prévu. C'est que bien des facteurs entrent en jeu, et l'équivalence catégorielle n'est que l'un d'entre eux, qui a trait au fonctionnement en discours. En d'autres termes, une catégorie correspond à un mode d'insertion et d'articulation dans le discours, et donc au paramètre pragmatique. Mais d'autres équivalences sont envisageables, et doivent même être envisagées.

3.3 D'AUTRES ÉQUIVALENCES

Un problème qu'on rencontre assez vite concerne ce que nous appellerons l'équivalence lexicologique. Elle a trait au problème du figement dans le monde parémique. D'une façon générale, les phrases sentencieuses ne sont pas figées, c'est-à-dire que le figement n'est pas un trait qui les caractérise¹⁷. On peut le voir à de nombreuses propriétés, la plus apparente étant l'existence, y compris pour un même locuteur, de variantes. Ainsi, la très connue forme *À cheval donné*

15 Litt. « La thune, c'est la thune ».

16 Qui se moque des Auvergnats, supposés avarés, et prononcer le *s* français comme un *ch*.

17 Sur ce sujet, v. Anscombe 20001, 2003, 2006.

on ne regarde pas la bride, admet entre autres variantes *À cheval donné, ne lui regarde pas en la bouche* (Dournon) ; *À cheval donné, on ne regarde pas à la bouche* (DesRuisseaux, Maloux, TLF) ; *À cheval donné, on ne regarde pas à la bride* (TLF, Gd Robert) ; *À cheval donné, on ne regarde pas la bouche* (Dela-court) ; *À cheval donné, on ne regarde pas les dents* (Lis et Barbier), etc. Ce qui n'empêche pas qu'on rencontre dans la catégorie des phrases sentencieuses – comme dans toutes les catégories – des éléments qui sont figés. Il semble par exemple n'y avoir aucune variante répertoriée à *À cœur vaillant, rien d'impossible*. Un traducteur peut donc parfaitement estimer qu'une traduction soignée doit fournir pour une phrase sentencieuse figée une phrase également figée, et conserver ainsi le niveau lexicologique. Cette louable intention ne dépasse malheureusement jamais le niveau du vœu pieux, pour des raisons qui ne dépendent pas de la volonté du traducteur. L'une d'entre elles provient des influences constantes et réciproques entre la langue parlée et les écrits littéraires. Dans le cas du français, un exemple typique est celui des *Fables* de La Fontaine. La Fontaine n'a pas inventé les morales de ses fables, et n'a fait que reprendre des expressions populaires, ou même littéraires, de son propre aveu. Ainsi la célèbre *La raison du plus fort est toujours la meilleure* se trouve dès le XVI^e siècle sous la forme *Où force règne, raison n'a lieu*, adage très vraisemblablement adapté du latin médiéval *Tunc jus calcatur, violentia cum dominatur*¹⁸. Le succès rencontré par les *Fables* a fait que la langue parlée a recueilli la morale de La Fontaine sans la modifier, ce qui est assez fréquent. L'équivalent espagnol *Donde fuerza viene, derecho se pierde*, bien qu'ayant également un auteur prestigieux¹⁹, ne bénéficie plus de cette aura, et a développé plusieurs variantes.

Par ailleurs, le choix de rendre une forme figée par une forme figée peut se heurter à d'autres exigences d'équivalence, en particulier l'équivalence statistique et l'équivalence stylistique. Chaque langue ayant son propre génie, ce qui est courant dans l'une peut être rare voire inexistant dans l'autre, et vice-versa. Considérons par exemple *Il n'y a pas de fumée sans feu*. Certes, l'espagnol possède la phrase parémique *No hay humo sin fuego* : mais alors que l'original français est d'un emploi banal et courant, son analogue espagnol l'est beaucoup moins, et « sonne » un peu livresque. On utiliserait beaucoup plus spontanément pour traduire l'idée recherchée *Cuando el río suena, agua lleva*²⁰. Choisir cette dernière expression serait donc faire prévaloir l'équivalence statistique. Autre exemple : l'espagnol *En la variedad está el gusto*, d'un usage très banal. Il n'y a en fait rien d'équivalent en français, si ce n'est *L'ennui naquit un jour de l'uniformité*, d'une utilisation nettement plus limitée que son analogue espagnol. Si l'on veut respecter l'équivalence statistique, il faudra se rabattre sur des tournures au caractère parémique peut-être moins marqué, mais statistiquement comparables à la leçon espagnole. Par exemple *Il faut savoir varier les plaisirs*, dont le statut de phrase parémique n'est cependant pas assuré : certains sujets parlants hésitent devant la combinaison *Comme on dit, il faut savoir varier les plaisirs*.

18 « Le droit est foulé au pied quand domine la force. »

19 À savoir le Marqués de Santillana, *Refranes que dicen las viejas tras el fuego*, 1454.

20 Litt. « Quand la rivière fait du bruit, c'est qu'elle charrie de l'eau. »

Le cas de *L'ennui naquit un jour de l'uniformité* nous met sur la piste d'une autre équivalence, à savoir l'équivalence stylistique. Cette phrase sentencieuse montre en effet, outre un vocabulaire assez « culte », un usage du passé simple tout à fait exceptionnel, et qui a été presque totalement remplacé par le présent²¹, hormis quelques rares survivants²². L'équivalence stylistique consistera donc à fournir une équivalence qui respecte le niveau stylistique, lequel se manifeste selon deux axes : un premier concerne le niveau de langue. Il peut être culte, neutre, familier, argotique, vulgaire, grossier, etc. Le problème réside en ce que chaque langue définit ces niveaux à l'intérieur de son système propre, et qu'il n'y a pas – ou pas nécessairement – correspondance entre niveaux stylistiques d'une langue à l'autre. Alors que l'espagnol admet sans problème d'utiliser des mots grossiers, le français est plus réticent sur ce point, et se contentera sans doute du niveau familier. L'argot est plus facilement employé dans le langage de tous les jours en français qu'en espagnol, où la séparation entre les deux niveaux est plus nette. D'où une recherche de correspondance de niveaux qui dans le cas du style est assez souvent délicate. En voici un exemple : la phrase sentencieuse *Quien quiera peces, que se moje el culo*, litt. « Qui veut des poissons, qu'il se mouille le c... ». Or si le mot *culo* est d'un usage banal en espagnol – il entre dans de nombreuses tournures familières – il est nettement plus marqué en français, où il est senti comme grossier et surtout vulgaire. Le traducteur s'efforcera donc de trouver un équivalent dans le style familier qui évite le piège de vouloir à tout prix utiliser l'équivalent lexical français de *culo*. Ce qui écarte d'entrée *Qui veut la fin veut les moyens*, qui d'une part, est d'un niveau stylistique très au-dessus du niveau familier requis, et d'autre part, fait table rase de la métaphore présente dans l'original.

Le second axe de l'équivalence stylistique concerne les variantes régionales. Certaines variantes sont en effet identifiées comme caractéristiques d'une région, ou simplement marquées à tort ou à raison comme « régionales ». Variation qu'on peut utiliser pour rendre certaines « atmosphères ». Supposons par exemple un roman paysan espagnol où l'un des protagonistes déclare que *Entre col y col lechuga*²³ « Il faut savoir varier les plaisirs ». Bien entendu, si l'esprit de la traduction est correct, la forme en revanche détonne dans l'atmosphère paysanne supposée. Et ne parlons pas de *L'ennui naquit un jour de l'uniformité*, qui porterait à rire²⁴... Heureusement, la nature (linguistique) fait bien les choses : il existe en français un *Changement d'herbage réjouit les veaux*, qui respecte l'équivalence sémantique *in situ* de façon me semble-t-il plus satisfaisante. On voit ainsi apparaître une nouvelle équivalence, l'équivalence sémantique. Il ne s'agit évidemment pas que les deux expressions source et cible aient le même sens, cette exigence allant de soi. Il s'agit de l'idée que ce sens soit atteint si possible par le même procédé dans les deux langues : à une métaphore on tente-

21 Il s'agit là d'un fait général. Ainsi, la très utilisée *Quien fue a Sevilla, perdió su silla* (« Qui va à la chasse perd sa place ») cède de plus en plus de terrain devant la version au présent *Quien va a Sevilla, pierde su silla*.

22 Ainsi le peu courant *Jamais bon cheval ne devint rosse*, ou l'espagnol plus banal *Quien tuvo, retuvo*.

23 Litt. « Entre du chou et du chou, de la laitue. »

24 Ce genre d'erreur est fréquent dans les doublages de films. Ainsi, dans la version française du fameux western *La Diligence*, le héros joué par John Wayne est censé être un humble *ranchero* de l'Ouest. Lorsqu'on l'avertit que la police risque de l'arrêter, il déclare « Je n'en ai cure », ce qui peut faire sourire.

ra donc de faire correspondre la même métaphore, ou du moins une métaphore relevant du même champ, ou d'un champ jouant un rôle analogue. Dans le domaine indo-européen contemporain, marqué par une culture de base commune, transmise par le latin et contrôlée par une religion (à peu près) commune, la tâche n'est pas surhumaine. Le fonds métaphorique est fréquemment partagé, ce qui ne se produit plus dans des cultures ayant d'autres origines. En particulier pour ce qui est des métaphores animales : elles sont typiques d'une culture, et un proverbe comme *Corneille dans la main vaut mieux qu'aigle en l'air* ²⁵, malgré une évidente parenté avec *Más vale pájaro en mano que ciento volando* et *A bird in the hand is better than two in the bush*, ne passera jamais pour un proverbe indo-européen moderne. Ni la corneille, ni l'aigle ne sont des familiers de notre parc animalier parémique. Cette équivalence sémantique fera préférer *Más vale un toma que dos te daré* à *Más vale pájaro en mano que ciento volando* comme traduction de *Un tiens vaut mieux que deux tu l'auras*, sur la base des métaphores présentes, bien que la seconde parémie espagnole soit plus courante que la première, et qu'on ne satisfait donc pas le critère d'équivalence statistique. Pour des raisons identiques, on pourra penser à traduire *Il ne faut pas vendre la peau de l'ours avant de l'avoir tué* par *No hay que vender la piel del oso antes de haberlo cazado*, plutôt que par *No hay que comerse el conejo antes de haberlo cazado*, fort peu utilisé. Cette seconde parémie étant de toutes façons beaucoup moins courante que la première, on satisfait de plus – au moins partiellement – à l'équivalence statistique.

Nous allons maintenant aborder le délicat problème de l'équivalence rythmique. L'observation et l'analyse montrent qu'en fait le trait parémique le plus saillant est la présence d'une structure rythmique. En d'autres termes, toute langue possède, prêts à l'emploi, un certain nombre de patrons rythmiques qui servent à construire des phrases sentencieuses. Ils sont en petit nombre, et se retrouvent également dans les comptines, les slogans, et partiellement dans les structures onomatopéiques. Ils correspondent à des moules poétiques caractéristiques d'une langue donnée, et la poésie habituelle n'a fait qu'utiliser ces moules en les épurant ²⁶. Ainsi, la phrase sentencieuse *Bonne renommée vaut mieux que ceinture dorée* peut se décomposer en *Bonne renommée / vaut mieux que / ceinture dorée*, c'est-à-dire un tercet de type a(4) b(3) a(4) ²⁷, avec de plus une structure rythmique de type /- -'/-'/- -'/, parfaitement symétrique. L'espagnol *Aunque la mona se vista de seda, mona se queda* est également un tercet a(5) b(6) b(5), de structure rythmique /- -'/-'/- -'/-'/- -'/, schéma également symétrique. Bien entendu, il s'agit là de cas d'école, il y a des cas moins réguliers. L'importance de ces structures rythmiques se manifeste de différentes façons. D'une part, les sujets parlants classent plus volontiers dans les proverbes les phrases sentencieuses ayant une structure rythmique que celles n'en ayant pas. Ils en arrivent même à identifier comme « proverbes » des phrases sentencieuses inventées de toutes pièces, mais pourvues d'une structure rythmique de bon aloi ²⁸. D'autre part, du point de vue diachronique, les formes sentencieuses

25 Recensé comme proverbe turc par Montreynaud, Pierron & Suzzoni.

26 Sur ces questions, v. Anscombre 2000a, 2000b, pour un aperçu rapide et non technique.

27 La lettre indique le type de la rime, et le chiffre le nombre de syllabes.

28 Si j'en crois les diverses expériences que j'ai menées sur le sujet.

empruntées à un auteur sont souvent remodelées pour s'ajuster à un patron métrique. Ainsi, *La critique est facile, mais l'art est difficile*, distique a(6) a(6), a été refait avec rime à partir d'un original *La critique est aisée, et l'art est difficile*, sans rime. L'espagnol *Perro ladrador, poco mordedor*, distique également, de structure a(6) a(6), est passé par différentes étapes antérieures, entre autres *Perro ladrador, nunca muerde*, a(6) b(4), et *Perro ladrador, nunca buen mordedor*, a(6) a(7).

On imagine aisément quel va être le problème. Si le traducteur désire conserver ce petit air authentique que confère la présence d'une structure rythmique, il lui faudra chercher une traduction possédant également une structure rythmique, et conforme aux schémas de la langue-cible, qui n'ont aucune raison d'être identiques à ceux de la langue-source, et encore moins d'apparaître simultanément ou non dans des formes sentencieuses équivalentes. On aura une idée de la difficulté dans le cas du passage de l'espagnol au français et vice-versa, si l'on considère que l'apparition de l'alexandrin dans la poésie française a eu une influence considérable au niveau des patrons rythmiques du monde parémique, phénomène qui ne s'est pas produit pour l'espagnol, qui montre donc une plus grande variété dans les rythmes.

4. LA TOUR DE BABEL : QUELQUES CAS ÉPINEUX

Comme on l'imaginera sans peine, le travail du traducteur de phrases sentencieuses n'est pas de tout repos, qui doit tenter de concilier les exigences des différentes équivalences que nous venons de passer en revue. Certes, il y a des cas idylliques : ainsi, à l'espagnol *De tal palo, tal astilla*, métaphorique et de structure rythmique a(4) b(4), correspond un *Tel père, tel fils*, également métaphorique, distique également de forme a(2) a(2), les deux phrases espagnole et française étant toutes deux fort courantes, et du même niveau stylistique neutre habituel. Mais la plupart du temps, le traducteur devra dans le cas qui nous occupe se contenter d'une cote mal taillée, voire même pire : on ne peut pas contenter tout le monde et son père !

Nous terminerons donc cet exposé par l'examen de quelques cas réels spécialement ardu, qui montreront qu'en toute dernière extrémité, c'est la sensibilité linguistique du traducteur plutôt que ses aptitudes techniques à la traduction qui fourniront une possible équivalence. De tels cas ne sont nullement l'exception, et comme nous allons le voir, la difficulté ne réside pas dans l'existence ou non d'une correspondance lors du passage d'une langue à l'autre, mais dans le fait que très souvent cette correspondance n'est pas une « bonne » correspondance, et foule au pied l'ensemble des équivalences que j'ai cru pouvoir dégager plus haut.

Commençons par le très connu *Cuando las barbas de tu vecino veas pelar, echa las tuyas a remojar*, litt. « Quand tu verras qu'on rase la barbe de ton voisin, mets la tienne à tremper », forme sentencieuse conseillant de prendre ses précautions quand le danger s'approche. Son statut sentencieux ne fait aucun doute, non plus que son caractère métaphorique. Par ailleurs, son patron rythmique est tout à fait remarquable : /*Cuando las barbas* / *de tu vecino* / *veas pelar* / *echa las tuyas* / *a remojar*/, c'est-à-dire a(5) b(5) a(5) c(5) a(5). C'est un quintil, et la disposition abaca est *grosso modo* celle du genre espagnol *romancillo*. Or il n'y a rien d'équivalent en français contemporain. La parémie parfois pro-

posée *Si on rosse ton voisin, prépare tes reins*, métaphorique et de structure a(6) a(4), est complètement sortie de l'usage contemporain et fait plutôt sourire. On n'a donc d'autre solution que de sortir du domaine restreint des proverbes, et de chercher un équivalent dans les phrases sentencieuses tout court. Bien entendu, cette solution du désespoir ne respecte à peu près aucune équivalence, si ce n'est la simple équivalence fonctionnelle. En fait, la valeur d'avertissement est assez bien rendue par la phrase situationnelle *Attention! Ça n'arrive pas qu'aux autres*, de création récente. Notons que le processus inverse se produit tout aussi bien : à une phrase-source sans caractéristique saillante correspond une phrase-cible d'une sous-catégorie bien particulière.

Considérons ainsi la phrase sentencieuse *La beauté ne se mange pas en salade* et tentons de lui trouver un équivalent en espagnol. Elle n'a pas de trait particulier, hormis la métaphore « alimentaire » attachée à *manger en salade*²⁹ : la seule beauté d'une femme ne fait pas bouillir la marmite. Il nous faut donc chercher une métaphore alimentaire attachée à la beauté féminine. La chance s'en mêlant parfois, l'espagnol possède la phrase parémique *Con hermosura sola, no se pone la olla*, qui correspond parfaitement pour le sens, présente également une métaphore alimentaire, est de même niveau stylistique, familier en l'occurrence. Mais elle possède de plus une structure rythmique de distique isosyllabique a(7) a(7), absente de l'original : c'est un proverbe. De ce point de vue, le traducteur auteur de *Sans fouet, enfant gâté* a correctement perçu dans l'original *Spare the rod, spoil the child* la présence d'un distique et a donné un distique inventé de structure a(3)b(4) comme traduction. Il aurait pu en fait aller plus loin en cherchant l'équivalent par le biais du fonds commun, à savoir *Qui parcit virgae, odit filium* (Bible, Proverbes). Or de nombreux auteurs³⁰ le donnent comme équivalant à *Qui bene amat, bene castigat*. Ce qui nous fournit immédiatement un *Qui aime bien, châtie bien*, distique de forme a(3) a(3).

Un sous-domaine sentencieux qui se révèle assez catastrophique pour la recherche d'équivalences, c'est celui des dictons agricoles et météorologiques. Les raisons en sont claires : d'une part les dictons agricoles et météorologiques s'appuient sur le calendrier liturgique, or les variations climatiques d'un pays à l'autre font que les mêmes événements ne coïncident pas nécessairement quant aux dates. D'autre part, et dans le cas qui nous occupe ici, il y a peu de similitudes entre les agricultures espagnole et française. Certes, il y a des éléments communs. Ainsi, *Santa Lucía, mengua la noche y crece el día* équivaut *grosso modo* à *À la saint Luce, le jour croît du saut d'une puce*, et *Lluvia suave quita vendaval* correspond à *Petite pluie abat grand vent*. Mais avec *Por San Blas / La cigüeña verás / Y si no la vieres / Año de nieves* lie la disparition ou continuation du froid au retour éventuel des cigognes pour la Saint Blaise (3 février). Corrélation ignorée du français *À la Saint-Blaise / L'hiver s'apaise / Mais s'il reprend / Longtemps on s'en ressent*. Que penser de *Hasta el cuarenta de Mayo, no te quites el sayo* qui recommande de se méfier des premiers beaux jours, alors que le français prétend à l'inverse qu'*En mai, fais ce qu'il te plaît* ! L'écart se creuse encore plus dans le domaine agricole, les cultures et l'élevage des deux pays

29 On peut cependant la découper en *la beauté / ne se mange pas / en salade*, ce qui en fait un tercet a(3) b(4) b(3), qui est une structure rythmique assez fréquente. De ce point de vue, ce serait donc un proverbe.

30 En particulier le très connu *Diccionario de expresiones y frases latinas*, de Victor-José Herrero Llorente, Madrid, Gredos, 2001.

respectifs présentant de grandes divergences. Là où le français affirme par exemple que *À la Sainte Catherine, tout bois prend racine* (25 novembre), l'espagnol ne dit rien. Voici un cas où l'on peut cependant s'en tirer. Un des derniers dictons encore présents dans le parler quotidien est *À la Saint Rémi, cueille tes fruits* (1^{er} octobre), distique a(5) a(4) recommandant donc la cueillette des fruits pour ce jour précis. L'espagnol voit plutôt le mois d'octobre en général comme le mois des semailles et surtout du retour du mauvais temps. À bien y regarder, il existe cependant des phrases sentencieuses relatives à la cueillette des fruits, à savoir : *A últimos de Noviembre, coge tu oliva siempre* (« À la fin novembre, toujours ramasser les olives ») ; *Agosto madura, y Septiembre vendimia la uva* (« Août mûrit et septembre vendange ») ; *Agosto tiene la culpa, Septiembre tiene la pulpa* (« Août a la coulpe, septembre a la pulpe ») ; *Por San Francisco, todo fruto es exquisito* (« À la Saint Francis, tout fruit est exquis »). Les deux premières ne conviennent guère : la première – distique a(8) a(8) – est relative au mois de novembre, et non au premier octobre, et concerne uniquement les olives. La seconde n'a trait qu'à la vendange, et parle de septembre. La troisième, également un distique a(8) a(8), dit que les fruits en général sont à point en septembre, mais ne parle pas du 1^{er} octobre. La dernière enfin, un tercet a(5) b(4) a(5), mentionne explicitement les fruits, et la Saint François a lieu le 4 octobre. Dans un contexte très général, c'est donc cette dernière qui conviendra le mieux, les autres pouvant par ailleurs convenir pour des correspondances plus spécifiques de date ou d'objet.

Nous voudrions terminer par l'exemple d'un proverbe espagnol très célèbre, pour lequel il n'existe rien d'équivalent en français contemporain, ce qui va nous obliger à adopter des solutions quelque peu désespérées. Il s'agit de *Más tiran dos tetas que cien carretas*, litt. « Plus tirent deux mamelles que cent charrettes ». D'utilisation courante, la leçon ci-dessus est un tercet / *Más tiran / dos tetas / que cien carretas* /, de structure a(3) b(3) b(5), qui possède de très nombreuses variantes, dont certaines fort lestes, et qui est très anciennement attesté³¹. Ce proverbe exprime dans un style familier le pouvoir de domination que le désir sexuel permet à la femme d'exercer sur l'homme. Or le français, malgré un grand nombre de phrases sentencieuses consacrées en bien ou en mal à la femme, ne possède rien qui ressemble à la phrase espagnole, même de loin. Il a pourtant existé en ancien français un *Plus tire cul que corde*³², proche de certaines variantes du proverbe espagnol³³, mais hélas, disparu aujourd'hui : l'idée – et l'image – n'est donc pas étrangère au champ parémique français. D'autant moins que La Curne dans son *Dictionnaire* cite un *Un cheveu de ce que l'on aime tire plus que quatre bœufs*, qui exprime une idée très proche. Pour nous tirer d'affaire, nous rappellerons une remarque faite plus haut : les sujets parlants identifient souvent un proverbe non à son contenu, mais à sa structure rythmique. Si la structure est adéquate, la forme est interprétée comme un proverbe d'une autre région, ou une variante dialectale. L'idée ici sera donc de forger de toutes pièces une phrase parémique correspondant à l'idée de la phrase espagnole, en lui conférant une structure rythmique qui la fera interpréter comme un authentique proverbe. Pour traduire *tetas*, mot familier mais non nécessai-

31 Dès le XVI^e siècle.

32 *Proverbes au vilain*, manuscrit, fin XIII^e.

33 En particulier *Más tiran nalgas en lecho que buyes en barbecho*.

rement vulgaire, nous utiliserons *tétons*, que nous ferons rimer avec *charreton*. L'idée est de reconstituer ainsi un tercet en respectant le maximum d'équivalences. La traduction de *Más tiran* par « tirent plus », comme dans la citation de La Curne est faible : l'espagnol *tirar* signifie ici « tirer de façon irrésistible, attirer fortement », ce pour quoi je suggérerai la traduction « tirer plus fort ». Ce qui nous donne la phrase parémique *Deux tétons tirent plus fort que cent charretons*, générique et métaphorique, et pourvue de plus d'un schéma rythmique / *Deux tétons / tirent plus fort / que cent charretons* / = a(3) b(4) a(5), très proche de l'original a(3) b(3) b(5). On peut bien entendu signaler le côté marginal de la traduction en l'assortissant d'un commentaire du style de *comme on dit chez moi / dans mon village*. On sauve ainsi l'esprit du texte parémique original – exigence minimale de toute traduction – mais on satisfait également certaines exigences propres au domaine des formes sapientiales.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ANSCOMBRE Jean-Claude, 1997, «Reflexiones críticas sobre la naturaleza y el funcionamiento de las paremias», *Paremia*, 6, p. 43-54.
- ANSCOMBRE Jean-Claude, 1999, «Estructura(s) métrica(s) en los refranes», *Paremia*, 8, p. 25-36.
- ANSCOMBRE Jean-Claude, 2000a, « Parole proverbiale et structures métriques », *Langages*, 139, p. 6-26.
- ANSCOMBRE Jean-Claude, 2000b, « Folklore, folklore. Est-ce que j'ai une gueule de folklore ? », *L'Observatoire de la Télévision*, 18 : 25-27.
- ANSCOMBRE Jean-Claude, 2001, «Refranes, polilexicalidad y expresiones fijas», *Actas del IVº Congreso Internacional de Lingüística Francesa*, Santiago de Compostela, p. 33-53.
- ANSCOMBRE Jean-Claude, 2003, « Les proverbes sont-ils des expressions figées ? », *Cahiers de lexicologie*, 1, p. 159-173.
- ANSCOMBRE Jean-Claude, 2006a, « Polyphonie et classification des énoncés sentencieux. Les marqueurs médiatifs génériques », *Le Français Moderne*, 74 (1), p. 87-99.
- ANSCOMBRE Jean-Claude, 2006b, « Les proverbes : un figement du deuxième type ? », *Linx*, 53, p. 17-33.
- AQUIEN Michèle, 1993, *Dictionnaire de poétique*, Paris, Le Livre de Poche.
- BARELLA Ana y CAMPOS Juana G., 1993, *Diccionario de refranes*, Madrid, Espasa Calpe.
- BERRENDONNER Alain, 1981, *Éléments de pragmatique linguistique*, Paris, Minuit.
- CERQUIGLINI Jacqueline et CERQUIGLINI Bernard, 1976, « L'écriture proverbiale », *Revue des sciences humaines*, 163, p. 359-375.
- COMBET Louis, 1996, «Los refranes : origen, función y futuro», *Paremia*, 5, p. 11-22.
- CONDE TARRÍO Germán, 1997, *Estudio comparativo de las paremias en francés, castellano y gallego*, Tesis Doctoral, Universidad de Santiago de Compostela.
- DELACOURT Frédéric, 1996, *Proverbes, dictons, et citations*, Paris, Éd. de Vecchi.

- DESRISSAUX Pierre, 1997, *Le Petit Proverbier : proverbes français, québécois et anglais*, Montréal, Bibliothèque Québécoise.
- DJAVADI Chafi, 1990, *Rouge du soir. Dictionnaire des dictons météorologiques*, Paris, Éd. Christian.
- DOURNON Jean-Yves, 1986, *Le Dictionnaire des proverbes et dictons de France*, Paris, Hachette.
- GARCÍA PEINADO Miguel Ángel, 1993, « Le fonds populaire chez La Fontaine : proverbes et locutions proverbiales », *Paremia*, 2 : 195-198.
- GIMENO Isabel, 1989, *El llibre dels refranys catalans*, Barcelona, Ed. Cap Roig.
- GONZÁLEZ José Luis, 1998, *Dichos y proverbios populares*, Madrid, Edimat.
- GROSS Gaston, 1996, *Les Expressions figées en français*, Paris, Ophrys.
- GUITER Henri, 1969, *Proverbes et dictons catalans*, Forcalquier, Éd. Robert Morel.
- JUNCEDA Luis, 1998, *Diccionario de refranes*, Madrid, Espasa Calpe.
- KLEIBER Georges, 1989, « Sur la définition du proverbe », *Recherches Germaniques*, 2, p. 232-252.
- LE ROUX DE LINCY Adrien-Jean-Victor, [1842] 1996, *Le Livre des proverbes français*, Paris, Hachette.
- LIS Michel et BARBIER Michel, 1980, *Dictionnaire du gai parler*, Paris, Éd. Mengès.
- MALOUX Maurice, [1980] 1995, *Dictionnaire des proverbes, sentences et maximes*, Paris, Larousse.
- MÉNDEZ PÉREZ Alejandra, 1996, « Mnemotecnica del refrán. La rima y las estructuras », *Paremia*, 5, p. 183-186.
- MEJRI Salah, 1996, *Le Figement lexical*, thèse d'État, Université de Tunis.
- MILNER George B., 1969, « De l'armature des locutions proverbiales. Essai de taxonomie sémantique », *L'Homme*, 9 (3), p. 49-70.
- MONTREYNAUD Florence, PIERRON Agnès et SUZZONI François, 1980, *Dictionnaire des proverbes et dictons*, Paris, Le Robert.
- MORAWSKI Joseph, 1925, *Proverbes français antérieurs au XV^e siècle*, Paris, Champion.
- NAVARRO DOMÍNGUEZ Fernando, 1993, « Hacia una nueva caracterización del concepto de paremia en su empleo lingüístico-discursivo », *Paremia*, 2, p. 21-26.
- PARÉS I PUNTAS Anna, 1997, *Diccionari de refranys Català-Castellà / Castellà-Català*, Barcelona, Edicions 62.
- PÉREZ MARTÍNEZ Héron, 1997, *Refrán viejo nunca miente*, México, El Colegio de Michoacán.
- PERMJAKOV Grigorij L'vovit, 1984, « Structural Typology of Paremias », *Kodikas / Code*, 7 (3/4), p. 263-268.
- PIERRON Agnès, 1997, *Dictionnaire des proverbes*, Allier (Belgique), Marabout.
- QUILIS Antonio, 1994, *Métrica española*, Barcelona, Ariel.
- REY Alain et CHANTREAU Sophie, 1997, *Dictionnaire d'expressions et locutions*, Paris, Le Robert.
- SEVILLA-MUÑOZ Julia, 1988, *Hacia una aproximación conceptual de las paremias francesas y españolas*, Madrid, Editorial Complutense.

SEVILLA-MUÑOZ Julia *et alii*, 2001, *1001 refranes españoles con su correspondencia en alemán, árabe, francés, inglés, italiano, polaco, provenzal y ruso*, Madrid, Ediciones Internacionales Universitarias.

SIMPSON John, 1982, *The Concise Oxford Dictionary of Proverbs*, London, Guild Publishing London.

TAMBA Irène, 2000, « Formules et dire proverbial », *Langages*, 139, p. 110-118.

LES FAUX AMIS, UNE QUESTION DE DEGRÉ : L'APPORT DE LA MÉTALANGUE SÉMANTIQUE NATURELLE

Bert PEETERS

Macquarie University, Sydney

Résumé — Il y a lieu de croire que derrière la grande majorité des faux amis, se cachent des similarités sémantiques qui risquent d'échapper à la vigilance du lexicographe, à moins qu'il ne se serve d'un outil descriptif aussi précis et puissant que la métalangue sémantique naturelle d'Anna Wierzbicka. Ladite métalangue s'utilise le plus souvent pour expliciter des différences que le recours à des termes trop complexes tend à masquer. Or, elle se prête également à la recherche de similarités dont la reconnaissance permet de mieux comprendre la nature du rapport qui existe entre des faux amis particuliers. Ceux-ci ne sont pas toujours « aussi faux », ni même toujours « aussi partiels », qu'il n'y paraît. Autrement dit, notre but est d'encourager une approche plus subtile des faux amis que celle dont on a l'habitude.

Mots clés — faux amis, expérience (*vs* experience), évidence (*vs* evidence), métalangue sémantique naturelle, primitifs sémantiques.

Abstract — It is not unreasonable to assume that large numbers of false cognates actually conceal semantic similarities which are likely to escape the notice of lexicographers, unless they adopt a descriptive tool as precise and as powerful as Anna Wierzbicka's Natural Semantic Metalanguage. The latter is typically used to explicate differences which tend to remain hidden if excessively technical jargon is relied on instead. Unexpectedly, perhaps, Wierzbicka's NSM can also be deployed in the search for similarities. Detecting such similarities allows us to better understand the nature of the relationship that exists between pairs of false cognates. These are not always "as false", nor are they always "as partial", as would appear to be the case at first glance. In other words, the aim of this paper is to call for a more subtle approach to false cognates than the one that has prevailed to date.

Keywords — false cognates, experience (*vs.* expérience), evidence (*vs.* évidence), Natural Semantic Metalanguage, semantic primes.

Il n'y a rien de particulièrement nouveau à l'idée selon laquelle on peut trouver, dans deux langues comparées à cet effet, des faux amis complets aussi bien que des faux amis partiels. D'après la définition d'Henriette Walter (2001 : 101), par exemple, les faux amis sont des éléments lexicaux « dont la forme, identique ou proche d'un mot de la langue cible, a en fait un sens complètement ou partiellement différent de celui qu'il a dans la langue de départ ». Pourquoi avoir cité cette définition-là plutôt que l'une des dizaines d'autres, plus ou moins analogues, qui ont été proposées dans la littérature ? Parce qu'à vrai dire, tout en paraissant parfaitement convaincante, elle présente un curieux contresens. Concrètement parlant, à en croire Walter, le mot français *seize* (l'un de ses propres exemples), d'orthographe identique à celle du mot *seize* en anglais (langue cible), aurait un sens différent de celui qu'il a... en français (langue de départ). Bref, il y aurait une différence sémantique entre les mots français *seize* et *seize* ! Corrigions donc la définition en disant que les faux amis sont des paires de mots de langues différentes et de sens partiellement ou totalement différents, en dépit d'une ressemblance aussi bien du point de vue de l'orthographe que de celui de la prononciation¹.

Voilà pour la définition. Reste à savoir s'il y a encore quelque chose de neuf à ajouter à un dossier déjà très volumineux et fort bien connu. Qu'y a-t-il encore à dire au sujet de l'existence de différents degrés de faux amis ? La thèse qui sera défendue dans ce qui suit est que la « métalangue sémantique naturelle » (*natural semantic metalanguage*) élaborée par Anna Wierzbicka et ses collègues permet de jeter une nouvelle lumière sur la question. À l'instar du maître de philosophie de Molière, nous dirons à ceux qui nous lisent : « Vous entendez cela, et vous connaissez la métalangue sémantique naturelle sans doute ? » – et aux Messieurs et Mesdames Jourdain qui nous répondront : « Oui, mais faites comme si je ne la connaissais pas. Expliquez-moi ce que cela veut dire », nous signalerons² que c'est une langue qui, comme toutes les langues naturelles, a un lexique et une grammaire. Seulement, ce lexique et cette grammaire sont d'un type particulier et font de la métalangue un outil descriptif d'une clarté maximale et d'une rigueur sans égal.

Le propre de la métalangue sémantique naturelle (dorénavant MSN), c'est qu'elle marie l'universalité de ses éléments constitutifs à la simplicité sémantique et grammaticale et à l'universalité de ses structures. Elle est donc culturellement neutre et on peut y recourir dans toutes les langues du monde. Il y a autant d'incarnations de la MSN, autant de MSN individuelles, en même temps distinctes et identiques, qu'il y a de langues naturelles. D'une part, chaque langue naturelle a effectivement sa propre MSN. Il y en a une pour l'anglais (développée essentiellement par Anna Wierzbicka et par son bras droit Cliff Goddard), une autre pour le français (développée par l'auteur de ces lignes, en collaboration avec Marie-Odile Junker), d'autres encore pour le swahili, le japo-

1 Du point de vue de l'orthographe, il y a souvent identité plutôt que simple ressemblance.

2 Pour plus de détails, voir les deux premiers chapitres du premier volume et les deux derniers chapitres du deuxième volume de Goddard & Wierzbicka (2002), le premier chapitre de Peeters (2006) et le premier chapitre de Goddard (2008). Pour des exposés récents en français, voir le numéro 54 de la revue *Linx* (Wierzbicka 2006a, b) et aussi Peeters (2010).

nais, le grec, etc. Chacune de ces MSN est en quelque sorte une forme simplifiée de la langue naturelle sur laquelle elle se greffe et qui lui sert de support. Chacune consiste en un fragment de langue naturelle, une mini-langue, avec un mini-lexique et une mini-grammaire. D'autre part, le propre de toutes ces métalangues, c'est qu'elles sont en dernière analyse isomorphes : leurs lexiques et leurs grammaires se correspondent rigoureusement. Autrement dit, en parlant de « la » MSN au singulier, on parle en fait d'un ensemble de variantes notationnelles dont on peut se servir pour décrire non seulement les faux amis, mais tous les mots dans toutes les langues – et non seulement tous les mots dans toutes les langues, mais aussi d'autres faits linguistiques, qu'il s'agisse de structures grammaticales, de normes communicatives, de valeurs culturelles, et ainsi de suite. Pour les détails de la syntaxe de la MSN, qui précise comment ses éléments constitutifs s'enchainent, quels enchainements sont admis (car universels) et lesquels ne le sont pas, on verra en particulier Goddard & Wierzbicka (2002), ou, dans une perspective romane, Peeters (2006). La présentation qui suit s'occupe avant tout du côté lexical de la MSN.

« Tout auteur de dictionnaires bilingues est conscient du fait qu'il faut s'attendre à ce que, d'une langue à l'autre, il y ait des différences sémantiques même entre des unités lexicales sémantiquement apparentées. Un grand nombre de mots anglais, par exemple, n'ont pas d'équivalents précis en français, et vice versa. Or, s'il en est ainsi quand le nombre de langues rapprochées s'élève à deux, que dire du postulat qu'il existe des mots qui, polysémie mise à part, ont des contreparties exactes dans *toutes* les langues, sinon qu'il est complètement fantaisiste ? » À cette question, formulée par Anna Wierzbicka dans le premier de deux cours magistraux délivrés au Collège de France en février 2004³, elle a elle-même apporté une réponse qui ne manquera pas d'étonner d'aucuns : à savoir que, loin d'être « complètement fantaisiste », la teneur générale du postulat a été confirmée à la suite de vérifications approfondies effectuées principalement au cours des vingt dernières années et portant sur des langues typologiquement très différentes (v. Goddard & Wierzbicka 1994 et 2002, Peeters 2006, Goddard 2008). Qui plus est, l'inventaire de ces mots – auxquels, dans le cas de la plupart des langues, il convient d'ajouter des syntagmes et/ou des affixes, dans la mesure où les mots d'une langue correspondent parfois à des syntagmes ou à des affixes dans une autre – a continué à être mis au point et comprend à l'heure actuelle une bonne soixantaine d'éléments appelés en français *primitifs sémantiques* (d'après l'anglais *semantic primes*). L'hypothèse originale, à savoir que le nombre d'éléments à retenir ne se chiffre ni par milliers, ni même par centaines, n'a pas été ébranlée, et on ne s'attend pas à ce qu'elle le soit. Cela ne veut pas dire que la liste est désormais définitive. Quoiqu'elle soit le résultat de longues années de recherches, son contenu reste susceptible de modification(s). Le tableau ci-après, organisé en rubriques, reflète l'état le plus récent du lexique de la MSN, version française⁴.

3 Le texte de ces cours a été publié dans le numéro 54 de la revue *Linx* (v. note 2).

4 Il y a quelques modifications par rapport à la liste fournie dans Peeters (2006), car la mise au point continue – et il s'ensuit, à plus forte raison, que les lexicalisations proposées dans des publications telles que Peeters (1994) et (1997) ne sont plus toujours représentatives de l'état actuel de la recherche. La convention la plus souvent adoptée au sein de l'approche MSN (traduction de l'anglais *NSM approach*) est d'imprimer les primitifs et leurs combinaisons au niveau infra-propositionnel en petites capitales, et les propositions complètes en caractères ordinaires.

JE, VOUS ⁵ , QUELQU'UN, QUELQUE CHOSE, GENS, CORPS	Substantifs
TYPE, PARTIE	Substantifs relationnels
CE, MÊME, AUTRE	Déterminants
UN, DEUX, BEAUCOUP, CERTAINS ⁶ , TOUT	Quantificateurs
BIEN, MAL	Évaluateurs
GRAND, PETIT	Descripteurs
PENSER, SAVOIR, VOULOIR, SENTIR, VOIR, ENTENDRE	Prédicats mentaux
DIRE, MOTS, VRAI	Discours
FAIRE, ARRIVER, BOUGER, TOUCHER	Actions, évènements, mouvement, contact
ÊTRE (QUELQUE PART), IL Y A, AVOIR, ÊTRE (QUELQU'UN / QUELQUE CHOSE)	Emplacement, existence, possession, spécification
VIVRE, MOURIR	Vie et mort
QUAND, MAINTENANT, AVANT, APRÈS, LONGTEMPS, PEU DE TEMPS, POUR QUELQUE TEMPS ⁷ , MOMENT	Temps
OÙ, ICI, AU-DESSUS, AU-DESSOUS, LOIN, PRÈS, CÔTÉ, DANS	Espace
NE... PAS, PEUT-ÊTRE, POUVOIR, À CAUSE DE, SI	Concepts logiques
TRÈS, PLUS	Intensificateur et augmentateur
COMME	Similarité

D'autres auteurs ont proposé des listes d'éléments soi-disant primitifs qui diffèrent de la liste ci-dessus dans la mesure où ce qu'on y trouve sont des éléments abstraits présumés universels d'un point de vue conceptuel, mais qui n'existent dans aucune langue particulière, ou bien des éléments qui ont l'air de coïncider avec des mots anglais (ou bien des mots d'une autre langue), mais qui n'en ont pas tout à fait le sens. Le lexique de la MSN, par contre, se compose d'éléments qui ont le même sens intuitivement intelligible pour tout le monde, et qu'il est dès lors impossible de comprendre de travers. On objectera peut-être que, dans le tableau ci-dessus, il y a des éléments polysémiques auxquels l'hypothèse d'un même sens intuitivement intelligible pour tout le monde ne s'applique pas. À cet égard, il importe d'insister sur le fait que l'identité de sens ne doit pas s'étendre – et le plus souvent ne s'étendra pas – à l'ensemble des contextes où les éléments lexicaux qui font partie du tableau sont susceptibles d'apparaître. Un mot polysémique ne pourra contribuer à la définition de ce qui est obscur, complexe, ou bien culturellement déterminé que dans un nombre strictement limité de ses sens. En principe, on n'en admettra qu'un seul, mais il faut bien se dire que c'est une règle qui n'est pas absolue (même si elle paraît valable au sein de la MSN française).

5 Plutôt que TU, proposé dans Peeters (2006). Pour une justification, voir Peeters (2009).
6 Peeters (2006) proposait IL Y A... QUI, comparable à la version italienne C'È... CHE. Si nous revenons aujourd'hui à une lexicalisation proposée dans des écrits antérieurs, c'est que la lexicalisation IL Y A... QUI ne se prête pas à la même combinatoire.
7 Plutôt que POUR UN TEMPS, proposé dans Peeters (2006), formulation qui paraît moins naturelle.

Les sens qu'on admet sont illustrés à l'aide d'une ou de plusieurs phrases dites « canoniques », qui illustrent l'usage qu'on peut faire des éléments lexicaux constitutifs du lexique de la métalangue. Voici quelques exemples de phrases canoniques (on parle aussi de *contextes* canoniques) ; elles se composent pour la plupart, mais pas exclusivement, de primitifs sémantiques :

Je ne veux pas que vous fassiez cela.

Quelqu'un m'a dit quelque chose, mais je ne l'ai pas entendu.

Ces gens ont vécu longtemps.

Cette chose a deux parties.

Si vous faites cela, les gens diront du mal de vous.

Quand je pense à vous, je sens quelque chose de bien.

Quand vous êtes avec moi, il ne peut m'arriver rien de mal.

Beaucoup de gens ne savent pas les mêmes choses que moi.

Il est arrivé quelque chose de bien dans cet endroit.

Une deuxième fonction des phrases canoniques est de préciser les propriétés grammaticales ou syntaxiques de chacun des primitifs. Rappelons que ceux-ci, loin de constituer simplement un lexique universel, se comportent comme des éléments lexicaux ordinaires : au sein des formules sémantiques où ils apparaissent, ils s'enchaînent à l'aide d'une syntaxe devenue au cours des années de plus en plus transparente, de plus en plus précise, *et* de plus en plus universelle. C'est d'ailleurs la raison pour laquelle un certain nombre d'entre eux possèdent un ou plusieurs *allolexes*, c'est-à-dire des variantes en distribution complémentaire, ce qu'illustrent à nouveau les phrases canoniques. Ainsi, ME et MOI sont des allolexes du primitif JE, et CHOSE est un allolexe de QUELQUE CHOSE. ENDROIT, quant à lui, est un allolexe du primitif OÙ, utilisé quand celui-ci se combine avec un déterminant ou un quantificateur (CET ENDROIT, UN AUTRE ENDROIT, BEAUCOUP D'ENDROITS).

D'aucuns s'étonneront que la liste des primitifs soit tellement limitée. À vrai dire, au départ, elle l'était bien plus encore : dans Wierzbicka (1972), ouvrage traduit du polonais, l'inventaire ne comportait qu'une douzaine d'éléments. C'est qu'il y avait – et il y a toujours – un autre critère d'inclusion (ou d'exclusion), qui s'ajoute à celui de l'intelligibilité intuitive pour l'ensemble des locuteurs de la langue naturelle correspondante. Il ne suffit pas qu'un élément lexical ait le même sens pour tout un chacun et que toute méprise sur ce qu'il signifie soit exclue ; il faut en outre que, dans l'emploi qu'on en fera dans l'une ou l'autre variante de la MSN, il soit sémantiquement simple, c'est-à-dire indivisible et indéfinissable. En effet, ce qui est sémantiquement divisible et dès lors définissable peut être divisé et dès lors défini de plusieurs façons (selon celui qui parle, selon les circonstances, etc.), et ce qui est susceptible d'être divisé et d'être défini de plusieurs façons peut être incorrectement interprété. L'inclusion d'éléments potentiellement ambigus est susceptible de semer la confusion et ainsi d'entraver le progrès de nos connaissances, ce qu'il faut bien entendu s'efforcer de prévenir. D'où la nécessité de distinguer ce qui est définissable et divisible de ce qui est indéfinissable et indivisible – autrement dit de ce qui est « primitif ». Les primitifs sémantiques ont été appelés ainsi précisément parce qu'ils sont indéfinissables et indivisibles. Ils constituent ce qu'on pourrait appeler un « noyau sémantique irréductible ». Les définitions

proposées au départ pour un certain nombre d'éléments qui, à l'heure actuelle, font partie de la liste se sont révélées inappropriées et ont été abandonnées. C'est ainsi qu'on en est arrivé, chemin faisant, à l'inventaire actuel. S'il s'avère, au bout d'un examen sérieux et exhaustif, qu'un élément qui avait été inclus dans le noyau sémantique irréductible est définissable, ou non universel, il faudra l'éliminer. Plusieurs arguments de ce genre ont été formulés mais se sont en général révélés non fondés. Tout comme les arguments qui ont été avancés en faveur d'éléments soi-disant injustement exclus de la liste, ou bien contre l'idée même d'un noyau sémantique irréductible et universel, ce sont pour la plupart des arguments non recevables qui reposent sur une évaluation inadéquate, trop superficielle, voire carrément fausse, de l'entreprise wierzbickienne.

Les lexicographes en particulier feront remarquer qu'il n'y a pas de mots indéfinissables, que tout peut être – et doit être – défini. Ils se trompent, et le résultat de leur jugement erroné est visible jusque dans les dictionnaires les plus respectés (à l'exception peut-être de certains dictionnaires pédagogiques écrits à l'intention des étudiants de langues étrangères). Définir, c'est rendre clair ce qui ne l'est pas. Cela présuppose que dans les définitions on n'utilise que des ressources qui sont sémantiquement plus simples que le mot à définir. Une définition qui est aussi obscure, voire plus obscure, que ce qu'elle cherche à définir n'est pas digne de ce nom ; ce n'est pas une définition au sens propre de ce terme. Définir quelque chose, c'est apporter des clarifications. Or, la réalité lexicographique est souvent très différente. Prenons, à titre d'exemple, la « définition » du mot *savoir* dans le *Petit Robert* : « appréhender par l'esprit ». Au lieu de l'éclairer, on a fini par obscurcir le sens du mot : des mots tels que *appréhender* et *esprit* sont sémantiquement plus complexes et infiniment moins neutres du point de vue culturel que le mot *savoir* qu'il s'agissait de définir (mais qu'on ne saurait définir dans la mesure où on a affaire à un primitif sémantique, c'est-à-dire à un élément lexical indéfinissable).

Définir, ce n'est pas seulement éviter la complexité. Il s'agit aussi d'éviter la circularité. Là encore, les lexicographes nous laissent le plus souvent sur notre faim. Un exemple de définition circulaire est celui que donne le *Lexis* des mots *nourriture* (« aliment destiné à entretenir la vie ») et *aliment* (« tout ce qui peut être digéré et servir de nourriture aux hommes et aux animaux »). Les « définitions » proposées ne nous informent guère quant à la différence précise entre ces deux mots : dans ce sens, elles laissent beaucoup à désirer, et on pourrait même faire valoir qu'il ne s'agit pas vraiment de définitions.

Il sera utile de récapituler ce qui précède. Wierzbicka et ses collègues ne se contentent pas de postuler, dans toutes les langues du monde, à côté des mots qu'il est possible de définir de façon claire et non circulaire, un noyau sémantique qui, lui, est irréductible ; ce serait, somme toute, un postulat relativement faible. L'hypothèse est bien plus radicale. En dépit du fait que ce sont, dans la plupart des textes publiés, des mots ou des syntagmes anglais qui leur servent de support formel, les primitifs sémantiques sont des universaux du lexique. Autrement dit, *le même noyau sémantique irréductible existe dans toutes les langues du monde*. Certes, il a fallu attendre les années quatre-vingt-dix pour que la « traduisibilité » des primitifs wierzbickiens (c'est-à-dire leur présence indubitable ou lexicalisation dans toutes les langues du monde) fasse l'objet de dépouillements systématiques allant au-delà des quatre langues (l'anglais, le

polonais, le russe et le français) que Wierzbicka avait utilisées jusque-là dans ses publications. Cependant, l'hypothèse à l'heure actuelle est que ce qui est indivisible, indéfinissable et intuitivement intelligible dans une langue quelconque l'est aussi dans toutes les autres : les primitifs sémantiques de l'anglais sont également ceux du français, du russe, du coréen, du nahuatl, etc. Les noyaux sémantiques irréductibles de ces diverses langues (et de toutes les autres) se correspondent rigoureusement ; il y a une isomorphie totale. Il est vrai que tout élément qui fait partie d'une langue particulière appartient à un réseau unique et autonome d'éléments et occupe une place particulière dans un réseau unique et autonome de relations. Dans ce sens, tous les mots de toutes les langues sont uniques et aucun ne devrait avoir des équivalents précis dans les autres langues du monde. L'idée qu'il puisse y avoir certains éléments linguistiques qui soient universels dans la mesure où ils possèdent un équivalent dans toutes les langues du monde est bien entendu d'autant plus extravagante. On trouve cependant des correspondances sémantiques, et ce sont précisément ces correspondances sémantiques qui, considérées dans leur ensemble, constituent le noyau sémantique irréductible que partagent toutes les langues du monde. Peu importe que les mots, les affixes et les syntagmes qui servent de support à la lexicalisation des primitifs sémantiques soient eux-mêmes polysémiques ; car, nous l'avons vu, il est indéniable que, très souvent, ils le sont. Peu importe que les polysémies observées puissent différer d'une langue à l'autre, et que certaines lexicalisations soient morphologiquement complexes (c'est-à-dire formellement transparentes). Les primitifs eux-mêmes sont monosémiques et sémantiquement simples, même si leurs lexicalisations ne le sont pas, ou ne le sont pas d'une façon systématique. Le plus souvent – et dans une langue telle que le français, il n'existe à ce principe aucune exception –, il n'y aura qu'un seul sens pour chaque lexicalisation qui soit en fait admis dans le noyau sémantique irréductible. Des contextes canoniques contribuent à élucider quel est le sens (ou quels sont les sens) qui sera (ou qui seront) admis, et quels autres sens seront par conséquent bannis.

Le moment est venu de passer au sujet proprement dit du présent exposé : à savoir la problématique des « faux amis », complets aussi bien que partiels. Le recours à la MSN, c'est-à-dire à un vocabulaire et à une grammaire sciemment postulés comme étant universels, présente non seulement l'avantage de faciliter la comparaison objective des langues et de leurs éléments constitutifs, sans privilégier l'une des langues en cause, c'est-à-dire sans tomber dans le piège de l'ethnocentrisme, ce qui est aussi important dans l'étude des faux amis que partout ailleurs. Il fournit également le moyen d'illustrer la thèse que, le plus souvent (gare aux généralisations hâtives !), les faux amis dits « complets » ne sont pas forcément « aussi faux » qu'on ne le croyait, ni les faux amis dits « partiels » « aussi partiels ».

Soyons plus précis. Des comparaisons corroborées à l'aide d'exemples concrets, comme on en trouve dans Van Roey *et al.* (1995), sont certes extrêmement utiles, et répondent à un besoin que jusque-là peu de dictionnaires de faux amis avaient reconnu (v. Granger & Swallow 1988 : 119). Cependant, elles ne vont pas assez loin. Grâce au recours à la MSN, plutôt que de dire que X_1 et X_2 sont des faux amis parce que celui-ci signifie A alors que celui-là signifie B (cas des faux amis complets), on pourra attirer l'attention sur d'éventuels points communs qui existent entre A et B – ce qui revient à dire que

ces prétendus faux amis complets sont « moins faux » qu'il n'y paraît. De la même façon, plutôt que de dire que Y_1 et Y_2 sont des faux amis, parce que celui-ci signifie C et D, alors que celui-là, par contre, signifie D et E (cas des faux amis partiels), on pourra attirer l'attention non seulement sur le sens partagé (D), mais aussi sur les points communs qui existent éventuellement entre les sens non partagés (C et E) – ce qui revient à dire que ces faux amis partiels sont « moins partiels » qu'il n'y paraît.

Il est impossible, dans l'espace qui nous est imparti, de multiplier les exemples et d'entreprendre en même temps une analyse et une description relativement approfondies. Aussi nous contenterons-nous d'une démonstration reposant sur deux paires de faux amis, réservant pour l'avenir des démonstrations plus ambitieuses et invitant d'autres à suivre notre exemple⁸.

1. EXPERIENCE EN ANGLAIS – EXPÉRIENCE EN FRANÇAIS

La première paire de mots retenue afin d'exemplifier la thèse présentée ci-dessus illustre également la règle générale selon laquelle, de deux faux amis, l'un français, l'autre anglais, c'est le mot français qui couvre le plus grand nombre d'emplois (Granger & Swallow 1988 : 109). Elle permet en outre d'illustrer le phénomène du rapprochement des faux amis, phénomène observé souvent de nos jours et qui, depuis quelque temps, affecte le français beaucoup plus que l'anglais. En gros on peut dire que, dans le français du XXI^e siècle, le mot *expérience* a quatre sens distincts, alors qu'en anglais contemporain le mot *experience* n'en a que deux. Les deux sens du mot anglais existent également en français. Un sens qui existait en anglais, mais qui est totalement sorti de l'usage, alors qu'il reste bien vivant en français, est celui qu'à l'heure actuelle, en anglais, on exprime à l'aide des mots *experiment* ou *experimentation*. *Expérimentation* existe en français, mais **expérimenter* n'existe pas. On dit *expérience*, et le mot s'utilise aussi bien au singulier qu'au pluriel, ainsi que le montre l'extrait suivant d'un fait divers où les deux usages sont juxtaposés⁹ :

- (1) En effet, les expériences d'isolement total dans une grotte, réalisées à plusieurs reprises par le spéléologue Michel Siffre, ont montré que dans le noir absolu, la dérive du temps est beaucoup plus importante que celle qu'aurait subie le naufragé de Madiran. [...] « Il est vrai que le même type d'expérience tentée par d'autres personnes a montré que chaque individu a sa propre horloge interne », ajoute toutefois Eric Gilli, géologue à l'université Paris-VIII et passionné d'exploration spéléologique. (F, 24 janvier 2005)

Parmi les verbes qui accompagnent souvent ce premier sens du mot *expérience*, il y en a qui fournissent des indications précieuses pour l'explicitation en MSN. *Tenter* est un verbe volitif : si on tente de faire quelque chose, c'est qu'on

⁸ En réfléchissant aux mots concernés et à ce qu'ils signifient, nous avons beaucoup profité de deux manuscrits non publiés d'Anna Wierzbicka, remaniés et intégrés dans Wierzbicka (2010). Toujours est-il qu'aucun effort n'a été entrepris afin de comparer les descriptions de Wierzbicka aux nôtres, d'autant plus qu'elles n'existaient encore que sous forme manuscrite et feraient très probablement l'objet de révisions avant d'être publiées (ce qui s'est vérifié par la suite).

⁹ La plupart des exemples français cités dans le texte proviennent d'un corpus de faits divers publiés dans des journaux français et belges entre le 1^{er} mars 1999 et le 15 août 2005. Les journaux cités sont identifiés à l'aide de sigles (le plus souvent suivis d'une date) : A (*L'Alsace*), D (*La Dépêche du Midi*), DH (*La Dernière Heure*), F (*Le Figaro*), M (*Le Monde*), P (*Le Parisien*), SO (*Sud-Ouest*).

veut le faire. *Montrer que* renvoie à un savoir auquel on aboutit. D'où l'explicitation suivante :

[A] X tente une expérience =

X fait quelque chose

parce que X veut savoir ce qui arrivera s'il fait cette chose

X ne peut pas faire cela

si avant maintenant X n'a pas pensé à ce qu'il veut savoir.

Les expériences de ce type (autrement dit les tests scientifiques) sont souvent *menées* dans un *laboratoire*, et peuvent être *renouvelées* :

- (2) Puisqu'une tribune lui est offerte, le trio du micro va, sous couvert de courbes et d'expériences de laboratoire, livrer un réquisitoire. (F, 22 février 2005)
- (3) Il a obtenu le résultat recherché. Malheureusement, il n'a pu mener plus loin l'expérience. (L, 18 février 2005)
- (4) Façon de dire que si le procès aboutissait à un échec en terme d'organisation, il serait peu probable que la chancellerie renouvèle l'expérience. (M, 1^{er} février 2005)

Puisque *expérience*, au sens de « test scientifique », n'a pas (ou n'a plus) d'équivalent anglais, on peut dire que, s'il n'avait que ce sens-là, le mot français *expérience* et le mot anglais *experience* seraient des faux amis complets. La réalité est évidemment différente, puisqu'au sens prospectif du mot *expérience* s'oppose en français un sens rétrospectif, qui existe également en anglais et que les deux langues partagent depuis plusieurs siècles. Il n'existe qu'au singulier. C'est le sens qui, si les autres sens n'existaient pas (ou n'avaient pas existé), ferait de notre paire une paire de « vrais amis ». En français, on peut illustrer le sens rétrospectif de la façon suivante :

- (5) Pour se défendre, il avance sa bonne foi : « J'ai une expérience de plus de vingt ans, même si je n'ai pas de diplôme, et j'ai suivi les avis de deux avocats, l'un français, l'autre irlandais. » (SO, 30 mars 2005)
- (6) Lui, dans son CV, résumait d'un trait sa vie : « plus de trente ans d'expérience et de réussite ». (M, 11 juillet 2004)
- (7) La solide expérience professionnelle de M^e Garreta, ancien bâtonnier de l'Ordre des avocats de Pau, rompu aux affaires réputées difficiles, a probablement pesé lourd dans la décision du bâtonnier en exercice, M^e Robert Malterre, qui a désigné d'office son collègue pour assurer la défense de Romain Dupuy. (SO, 4 février 2005)
- (8) Cependant, on doit pour juger cet aspect des choses tenir compte aussi de l'expérience des personnes qui entendent l'épouse Fourniret. Ni le juge d'instruction Bernard Claude, ni le commissaire Jacques Fagnart ne sont tombés de la dernière pluie. (LB, 3 mars 2005)

Les deux premiers exemples fournissent une indication très concrète du fait que l'expérience dont il est ici question est un attribut normalement associé à un âge relativement avancé grâce auquel la personne qui a de l'expérience a pu accumuler de nombreux savoirs et de nombreuses aptitudes. Les deux autres exemples confirment cette impression par le biais de l'adjectif *solide* et de l'expression idiomatique *ne pas être tombé de la dernière pluie*, dont la connotation positive souligne en outre qu'il est bon d'avoir de l'expérience – un jugement que les premiers exemples reflètent d'une façon moins directe (à l'aide

de la proposition *pour se défendre* et de la juxtaposition des mots *expérience* et *réussite*). Soit, en MSN :

[B] X a de l'expérience =

certaines gens peuvent dire quelque chose comme ça :

beaucoup de choses me sont arrivées, pas au même moment

j'ai vu beaucoup de choses, pas au même moment

j'ai entendu beaucoup de choses, pas au même moment

j'ai senti beaucoup de choses, pas au même moment

j'ai fait beaucoup de choses, pas au même moment

j'ai souvent pensé à ces choses

à cause de cela, je sais beaucoup de choses maintenant

comme quelqu'un peut savoir beaucoup de choses s'il vit pendant longtemps

X est quelqu'un comme ça

c'est bien si quelqu'un est comme ça.

En anglais, on peut illustrer le sens rétrospectif (*X has experience*) de la façon suivante ¹⁰ :

- (9) With more than thirty years experience and more than thirty-five Colombian Emeralds International store locations throughout the Caribbean, we have built an outstanding reputation for integrity, quality and value.
(www.discountclub-aruba.com/index.asp?id=58&details=54)
- (10) The internships are developed to provide students with solid professional experience that they can use as a springboard into their career and we work with host organizations to assure that the experiences are appropriate for students.
(<http://ie3global.oregonstate.edu/ie3/advisors/advisorFAQ.html>)
- (11) This indicates that over time, and compared with qualifications at degree level and above, the labour market is valuing the experience of those who have studied at lower levels more than the qualification itself.
(<http://educationcounts.edcentre.govt.nz/indicators/edachievmnt/tspar33.html>)
- (12) It takes time to gain the necessary skills and experience to achieve success.
(COBUILD) ¹¹

Le troisième sens (le deuxième du mot anglais) n'est ni rétrospectif ni prospectif, mais introspectif. Il existe non seulement au singulier, mais aussi au pluriel, sans doute parce qu'il s'agit d'expériences comptables qui peuvent avoir une certaine durée, ou bien être momentanées. Dans la mesure où elles laissent une marque – v. aussi (26) ci-dessous –, on se rappelle les expériences dont il est ici question : on est conscient de l'expérience au moment même où elle se déroule, on éprouve à ce moment-là une ou plusieurs émotions, et c'est de l'expérience autant que des émotions qu'on se souvient. Les exemples (13) à (15) proviennent de notre corpus (v. note 9) ; les autres ont été repérés sur

¹⁰ Afin de souligner l'équivalence du sens rétrospectif dans les deux langues, la sélection des exemples a été en grande partie guidée par les extraits français (5) à (8). Toutes les pages web citées ont été consultées le 15 septembre 2006.

¹¹ Le corpus COBUILD (<http://titania.cobuild.collins.co.uk>) est la propriété intellectuelle de Collins. Les exemples retenus sont cités d'après les manuscrits mis à notre disposition par Wierzbicka (v. note 8).

internet. On notera dans (13) et (14) le recours à l'adjectif *première*, et dans tous les exemples la présence d'une composante émotive – implicite dans le cas de (15) où *expérience* équivaut à *expérience sexuelle*.

- (13) Pour Catherine Thévenon, psychiatre, « la rupture sentimentale prend d'autant plus d'importance chez les adolescents qu'elle constitue la première expérience soumettant à des émotions fortes alors que le jeune ne possède pas suffisamment de connaissances intellectuelles et cognitives pour réagir. » (F, 28 janvier 2005)
- (14) Cette première expérience en communauté, dans le cadre d'une classe de mer d'une semaine, lui est-elle parue trop dure, au point qu'il cherche à s'enfuir en pleine nuit ? (F, 10 avril 2004)
- (15) Sa déception fut cruelle quand sa femme, infirmière, lui révéla qu'elle avait déjà eu une expérience avant lui. (M, 11 juillet 2004 ; extrait d'un rapport psychiatrique)
- (16) Quand l'eau est bien claire, quoi de plus beau que de prendre des bouteilles d'oxygène et de plonger au milieu des poissons ? C'est une expérience fantastique, mais il faut une formation d'abord, car cette activité n'est pas sans dangers. (www.clg-camus-argenteuil.ac-versailles.fr/rss.php)
- (17) Comment peut-on trouver la force de lutter encore, contre toute raison, alors que la mort semble déjà avoir gagné ? C'est cette expérience effroyable que Joe Simpson raconte dans ce livre, devenu un bestseller international. (www.amazon.fr/Mort-suspendue-Joe-Simpson/dp/2723445798)
- (18) Vous prenez conscience que ces militaires qui font cette expérience horrible ne sont que de jeunes hommes – les simples soldats ayant 18 ou 19 ans, et les officiers pas plus de 24 ans. (<http://archives.arte-tv.com/fiction/warriors/ftext/apropos.htm>)
- (19) Adhéz à un stage de pilotage de F1, c'est génial et c'est une expérience inoubliable. (<http://www.aquadesign.be/news/article-2017.php>)

À ces exemples français, on peut ajouter les exemples anglais que voici :

- (20) This first experience proved to be much easier than expected. We did not expect to reach the Confluence on snowmobiles. Nevertheless, we are very happy and proud of our success. (www.confluence.org/confluence.php?lat=51&lon=-77)
- (21) Domestic/food prep in a 5-star hotel in Lake Louise, Alberta. It was a fantastic experience. (COBUILD)
- (22) The cows started shunting me against the post and I had to grab onto it to keep my feet. It was a terrifying experience as they pressed against me with their angry eyes glaring menacingly. (COBUILD)
- (23) Elma was so cold that she was unable to speak for almost half an hour. [...] It was a horrendous experience. (COBUILD)
- (24) An unforgettable experience is an event that has a great impact on a person's life. It can also be something that can change a life for ever. My unforgettable experience is when I became a Christian. (www.megaessays.com/viewpaper/12391.html)

En français, le sens introspectif est relativement nouveau. Il faudrait entreprendre des recherches plus approfondies pour savoir quand il a fait son apparition. Les différences entre deux éditions du *Robert & Collins* (1996 et 2002) sont symptomatiques. La plupart des expressions toutes faites continuent à être traduites à l'aide de mots autres que le mot *expérience*. D'autre part, les deux

éditions proposent de traduire *it was a new experience for me* en disant *c'était une nouveauté* ou *une expérience nouvelle pour moi*. *She swam in the nude and it was an agreeable experience*, traduit en 1996 par *Elle a nagé toute nue et a trouvé cela agréable*, est traduit en 2002 par *Elle a nagé toute nue et a trouvé l'expérience agréable*.

Il y a entre le sens introspectif du mot anglais et le sens introspectif du mot français une différence que la longue liste d'exemples allégués ne doit pas masquer. *Le Robert & Collins* (2002) décrit le sens français à l'aide des mots *aventure humaine* et offre, à titre d'illustration, le syntagme *expérience amoureuse* ou *sexuelle*, exemplifié de façon elliptique dans (15), et les expressions idiomatiques *tenter l'expérience*, *faire l'expérience de quelque chose* et *faire une expérience de quelque chose*. Si on ajoute que l'expérience traumatique de l'exemple (13) concerne la rupture d'une adolescente de quinze ans avec son petit ami, rupture suivie d'un suicide, et que l'expérience en communauté de l'exemple (14) est celle d'un séjour dans un centre de vacances¹², il est relativement clair qu'il semble y avoir en français une dynamique qui manque en anglais. L'accent est mis beaucoup plus sur ce qu'on a fait, ce qu'on fait ou ce qu'on fera que sur ce qui est arrivé, ce qui arrive ou ce qui arrivera à celui qui fait l'expérience (v. Van Roey *et al.* 1995 : 286). Certes, si quelque chose arrive à quelqu'un, c'est souvent à cause de ce que fait la personne en question. Mais ce n'est pas toujours le cas : la différence entre le français et l'anglais est indéniable. Il n'est pas sans intérêt de noter qu'en anglais on ne peut ni **do* ni **make an experience*, mais qu'on peut très bien *have an experience* (*faire et avoir une expérience* sont l'un et l'autre possibles en français).

Ce qui précède peut être explicité en MSN de la façon suivante :

- [C] c'est une expérience (inoubliable, horrible, curieuse...) =
 parfois, quelqu'un peut penser quelque chose comme ça :
 à ce moment-là, dans cet endroit-là, j'ai fait quelque chose
 je savais que je le faisais
 à cause de cela, je sentais quelque chose
 je savais que je me sentais comme ça
 quand ce quelqu'un pense à ce qu'il a fait, il peut y penser comme ça :
 je sais ce que j'ai fait
 je sais ce que je sentais
 à cause de cela, je veux dire quelque chose
 je pense comme ce quelqu'un
 à cause de cela, je dis quelque chose de ce que j'ai fait
 quand je dis cela, je sens quelque chose
 d'autres gens ne peuvent pas sentir ce que je sens
 s'ils n'ont pas fait la même chose
 d'autres gens ne peuvent pas savoir ce que je sens

¹² Les autres exemples français se passent de commentaires. Voir la note 13 pour une interprétation différente de l'exemple (14).

s'ils n'ont pas fait la même chose

[D] it's a(n) (unforgettable, horrific, curious...) experience =

parfois, quelqu'un peut penser quelque chose comme ça :

à ce moment-là, dans cet endroit-là, il m'est arrivé quelque chose

je savais qu'il m'arrivait quelque chose

à cause de cela, je sentais quelque chose

je savais que je me sentais comme ça

quand ce quelqu'un pense à ce qui lui est arrivé, il peut y penser comme ça :

je sais ce qui m'est arrivé

je sais ce que je sentais

à cause de cela, je veux dire quelque chose

je pense comme ce quelqu'un

à cause de cela, je dis quelque chose de ce qui m'est arrivé

quand je dis cela, je sens quelque chose

d'autres gens ne peuvent pas sentir ce que je sens

si la même chose ne leur est pas arrivée

d'autres gens ne peuvent pas savoir ce que je sens

si la même chose ne leur est pas arrivée.

Plutôt que d'accentuer la différence entre l'usage introspectif français et anglais, comme on le fait d'habitude dans les dictionnaires de faux amis, les explicitations en MSN insistent autant sur ce que les deux usages partagent que sur ce qui les sépare. Elles montrent que les éléments partagés sont nombreux, la seule différence, exprimée plusieurs fois, étant le recours à un primitif sémantique (FAIRE) plutôt qu'à un autre (ARRIVER) dont la syntaxe est différente (« X [FAIRE] Y » vs « Y [ARRIVER] à X »). On notera par ailleurs que la perspective adoptée est celle de la personne qui a eu ou qui a fait l'expérience. C'est qu'une expérience est un événement à répercussions personnelles : la personne qui a eu ou qui a fait l'expérience sait, mieux que quiconque, ce qu'elle a fait ou ce qui lui est arrivé, ce qu'elle a ressenti, et c'est donc elle qui est en mesure d'en parler avec le plus d'autorité. Ceux qui décrivent les expériences d'autrui – comme la psychiatre de l'exemple (13) ou le fait-diversier lui-même dans l'exemple (14) – se mettent pour ainsi dire (ou essaient avec plus ou moins de succès de se mettre) à la place de « l'expérienceur ».

Il a été signalé plus haut qu'il y a eu un rapprochement entre le français et l'anglais, à la suite duquel nos deux faux amis sont devenus « un peu moins faux ». Ce rapprochement est tout récent, et peut être illustré à l'aide des exemples ci-dessous :

- (25) Les deux rescapés feront part de leur expérience au comité d'hygiène et de sécurité (CHS). (F, 2 mars 2005)
- (26) Marquée à jamais par cette expérience, Valérie pense surtout à l'avenir de Marion. (P, 30 octobre 2004)
- (27) Attendre un bébé est une expérience merveilleuse pour tous les parents. (www.educh.ch/coaching-formation-new.php?New=44)

- (28) Louise Madeleine, qui voit ainsi partir à l'échafaud non seulement son grand-père mais également ses deux parents ainsi que le couple des Chateaubriand que formaient sa sœur aînée et son époux, ne se remettra jamais de cette expérience terrifiante qui l'a définitivement traumatisée alors qu'elle n'était âgée que de vingt-et-un ans.

(www.tocqueville.culture.fr/fr/portraits/p_famille-mere.html)

L'expérience dont les deux rescapés (25) feront part est celle de l'incendie dans le tunnel du Mont-Blanc en mars 1999. L'expérience qui a marqué Valérie (26) est d'avoir été impliquée malgré elle dans une tuerie qui venait de se dérouler chez les voisins et qui risquait de se poursuivre chez elle. Dans les deux cas, il est moins question de ce que les deux rescapés et de ce que Valérie ont fait que de quelque chose qui leur est arrivé ; et pareillement dans le cas des parents qui attendent un bébé (27) et de Louise Madeleine qui perd toute sa famille sous la guillotine (28)¹³. Le meilleur moyen de rendre compte du rapprochement entre le mot *expérience* en français et le mot *experience* en anglais est d'admettre un quatrième sens pour le mot français, observable notamment dans des énoncés du type *Il m'est arrivé une expérience (inoubliable, horrible, curieuse...)*¹⁴. Ce quatrième sens est proche du troisième, mais tout de même distinct – et correspond à celui que nous avons donné ci-dessus sous [D]. Si, en revanche, on optait pour la disparition du sens [C] en faveur du sens [D], importé de l'anglais, il serait difficile d'expliquer pourquoi les expressions verbales comportant le verbe *faire* sont – et restent – tout à fait naturelles en français, alors qu'elles sont inadmissibles en anglais.

2. EVIDENCE EN ANGLAIS – ÉVIDENCE EN FRANÇAIS

Passons à la deuxième paire, qui regroupe les mots *evidence* et *évidence*. “There is no word for ‘evidence’ in other European languages.” Voilà ce qu'observe, dictionnaires à l'appui, Wierzbicka (2002 : 95)¹⁵. Qu'en est-il alors du mot français *évidence*, qui figure dans le titre du célèbre livre de Raymonde Carroll (*Évidences invisibles : Américains et Français au quotidien*, 1987) ? L'ouvrage a été publié en anglais en 1988, avec pour titre principal *Cultural misunderstandings*, c'est-à-dire, en français, « malentendus culturels ». Et pourquoi pas « Invisible evidences » ? Grammaticalement parlant, parce que le pluriel *evidences* est sorti de l'usage de tous les jours et ne survit que dans les textes de l'apologétique chrétienne anglo-américaine ; sémantiquement parlant (et c'est bien plus important), parce que le mot n'a pas le même sens. *Évidence* et *evidence* sont des faux amis. Le mot adopté par la traductrice partout où Carroll recourt au mot *évidence* est *verity* – mot très peu usité en anglais, et qui rappelle le mot français *vérité*, sans en avoir tout à fait le même sens, *vérité* se traduisant en anglais à l'aide du mot *truth*. Autrement dit, *vérité* et *verity* sont eux aussi des faux amis. Ils ne seront pas abordés dans ce qui suit.

13 De même, on pourrait faire valoir, dans le cas de (14), que la disparition est la conséquence, non pas de ce que le jeune disparu a fait dans le centre de vacances, mais de ce qui lui est arrivé au cours de la semaine.

14 Les énoncés de ce type rencontrent encore une certaine résistance (il y a des locuteurs natifs qui les rejettent), mais ils se répandent.

15 Observation fondée sur la consultation de quatre dictionnaires bilingues (anglais / italien, anglais / français, anglais / allemand et anglais / polonais).

Parlant de sa culture de Française, qu'elle définit comme « la logique selon laquelle [elle] ordonne le monde » (1987 : 17), Carroll (*ibid.* : 18) explique ainsi le titre de son livre :

De cette logique, il est une part qui est tacite, qui est invisible, et c'est la plus importante. Il s'agit des prémisses dont nous tirons constamment nos conclusions. Ces prémisses, nous n'en avons pas conscience parce qu'elles sont, pour nous, des évidences. C'est tout ce qui, pour nous, « va de soi », et est donc transparent.

À deux reprises, l'auteur recourt au groupe prépositionnel *pour nous*. C'est un détail important. Ce qui est une évidence pour les uns n'est pas nécessairement une évidence pour les autres. Il y a toujours des gens (et parfois il y en a beaucoup) à qui l'une ou l'autre « évidence » échappe. On peut aller plus loin : ce qui est une évidence *pour moi* n'est pas nécessairement une évidence pour celui à qui je m'adresse (et pareillement pour les deux autres personnes verbales). Tout dépend de la nature de l'évidence. Cependant, cette restriction n'est pas prise en compte par celui qui parle. Dans notre corpus de faits divers se trouve l'extrait suivant :

- (29) « À partir de mardi, on va basculer sur l'enlèvement, *c'est une évidence* », a déclaré, dimanche, le lieutenant-colonel Stéphane Ottavi, commandant de la section de recherche de gendarmerie de Rennes. (M, 13 avril 2004)

Ce qui est une évidence pour le lieutenant-colonel Ottavi n'en est pas une pour tout le monde. Aussi la journaliste (Alice Coffin) se voit-elle obligée de clarifier ce qui, aux yeux de son interlocuteur, est une évidence – pour lui, et sans doute pour tous ceux qui sont au courant du système judiciaire français en matière de disparitions : à savoir que ce n'est qu'au bout d'une semaine qu'une disparition devient officiellement suspecte. Elle écrit – et il convient d'attirer l'attention sur le marqueur *en effet*, qui introduit la clarification qui suit :

- (30) En effet, mardi soir, la durée légale de l'enquête pour flagrance, ouverte mercredi 7 avril par le procureur de Saint-Nazaire, Jacques Noguellou, expirera. Le parquet de Saint-Nazaire devrait alors ouvrir une information judiciaire. (*ibid.*)

Dans un autre extrait se trouvent reproduits les propos d'un infirmier qui s'inquiète au sujet de l'avenir de sa profession :

- (31) Rendre la profession attractive, améliorer les conditions de travail et de sécurité, c'est une évidence. (SO, 2 février 2005)

Les propositions infinitives dans cet exemple sont elliptiques : ce qui est une évidence, non seulement pour l'infirmier qui parle mais aussi pour tous ceux qui sont au courant de l'état dans lequel se trouve le système de santé publique français, c'est qu'il est *indispensable* de rendre la profession attractive et d'améliorer les conditions de travail et de sécurité.

Un autre élément clé sur lequel insiste Carroll dans sa description de la part tacite d'une culture, c'est que nous ne sommes pas conscients des prémisses dont nous tirons nos conclusions *parce que ce sont des évidences*. Autrement dit, une évidence est un savoir qui s'impose à nous, qui va de soi. Est-ce à dire qu'une évidence est un savoir auquel on ne s'arrête jamais et qui n'a pas besoin d'être explicité (ou d'être rappelé) ? Bien sûr que non : les exemples cités en apportent la preuve. En principe, le mot *évidence* s'utilise au sujet d'un savoir qui vient d'être énoncé et dont, aux yeux du locuteur, et pour s'exprimer à l'aide

de primitifs sémantiques, « les gens ne peuvent pas ne pas savoir qu'elle est vraie ». C'est donc quelque chose qui est intuitivement vrai, dont la vérité ne demande pas à être établie par le biais de la pensée. Du coup, il est impossible de soumettre une évidence à la réflexion et d'aboutir à la conclusion qu'après tout elle manque de vérité. Si jamais on arrive à cette conclusion-là, c'est que la soi-disant évidence n'en était pas une. C'était une pseudo-évidence. En MSN, cela se laisse expliciter de la façon suivante (P, dans la formule ci-dessous, signifie « proposition ») :

[E] P, c'est une évidence =

je dis P

je sais que P est vrai

je pense :

les gens ne peuvent pas ne pas savoir que P est vrai

à cause de cela, quand les gens pensent à P, ils ne peuvent pas dire que P n'est pas vrai.

Deux remarques s'imposent. Dans la première des propositions qui suivent la principale « je pense », on trouve le verbe *pouvoir* assorti d'une négation et suivi d'un infinitif négatif. Dire qu'on *ne peut pas ne pas savoir* quelque chose, c'est faire allusion à la nature intuitive du savoir décrit comme une évidence. Quant au sujet du verbe *pouvoir*, il constitue une espèce de généralisation : il ne s'agit pas vraiment de « tout le monde » (ou de « tous les gens »). Peut-on se contenter d'un à-peu-près de ce genre ? À vrai dire, on n'a pas le choix. Il y a lieu de croire qu'il est impossible d'être plus précis. En disant, par exemple, à propos d'une évidence, que celui qui vient de l'énoncer pense que « beaucoup de gens » (plutôt que « les gens » tout court) « ne peuvent pas ne pas savoir que P est vrai », on invite automatiquement la réflexion « mais il y en a d'autres ». Or, le propre d'une évidence, c'est justement que cette réflexion ne s'opère pas d'une façon consciente. Autrement dit, la réflexion « qu'il y en a d'autres » est elle-même une évidence¹⁶.

Par ailleurs – c'est la deuxième remarque –, il convient de préciser que dans notre corpus de faits divers, sur 46 occurrences du mot *évidence* (dont 40 au singulier et 6 au pluriel), il n'y en a que deux qui, syntaxiquement, se rapprochent de la tournure « P, c'est une évidence ». Les autres occurrences (voir la liste distributionnelle en annexe) entrent pour la plupart dans des expressions idiomatiques (par exemple 12 fois *mettre en évidence*, 10 fois à *l'évidence*, 7 fois de *toute évidence* / *d'évidence*) qu'il faudrait expliciter séparément. D'une façon ou d'une autre, les composantes dégagées ci-dessus auront un rôle à jouer. Mettre quelque chose en évidence, par exemple, c'est faire de sorte que ce qui n'était pas encore perçu comme une évidence est perçu

16 La généralisation proposée ici est comparable à celle qui figure en tête de la majorité des scénarios culturels (ou « évidences culturelles ») proposés dans la littérature MSN. Que veut dire au juste la proposition « Les gens pensent comme ça », sur laquelle s'ouvrent ces scénarios ? Paraphrasant Wierzbicka (2006c : 34), on pourrait se demander, au sujet d'un scénario culturel explicitant une façon de penser associée aux Français : Qui sont-ils, les gens qui pensent « comme ça » ? La réponse, c'est que *beaucoup* de Français pensent « comme ça ». Il est évident que tous les Français ne pensent pas de la même façon ; cependant, même ceux qui ont des idées différentes savent que la plupart des Français pensent de la sorte, et qu'il est en principe impossible de se soustraire au scénario culturel en question sans se faire remarquer.

de cette façon-là à partir d'un certain moment. Et c'est à partir de ce moment-là que les gens *ne pourront pas ne pas savoir* que ce qui a été mis en évidence est vrai, et qu'ils *ne pourront pas dire*, en y réfléchissant, que ce qui vient d'être mis en évidence n'est pas vrai.

Passons à l'anglais *evidence* ; on le fera en évoquant d'abord quelques emplois français du mot *évidence* (le plus souvent au pluriel) qui semblent correspondre au mot anglais *evidence* (au singulier). Voici, pour commencer, les propos d'un enquêteur judiciaire dans l'affaire Fourniret (v. Peeters 2007) :

(32) Fourniret ne parle qu'une fois le dos au mur, devant les évidences. On va en chercher et il parlera. (DH, 13 juillet 2004 ; P, 14 juillet 2004)

Une traduction anglaise du genre "*Fourniret speaks only with his back to the wall, faced with the evidence. We'll look for it and he'll talk*" paraît tout à fait possible. Reste à savoir quelles sont les évidences auxquelles pense l'enquêteur. Le lecteur de langue anglaise risque d'arriver à la conclusion qu'il s'agit de vêtements (!). Il n'en voudra pour preuve que des indications contextuelles du type que fournit DH, qui poursuit :

(33) Apparemment, Fourniret aurait donc conservé de dangereuses preuves dans les affaires saisies. Des vêtements de ses victimes. (*ibid.*)

Il est indéniable que des vêtements peuvent être considérés comme *evidence*. Mais est-il possible d'y voir des *évidences* ? La réponse de la grande majorité des francophones sera sans doute négative : les évidences ne sont pas les vêtements à l'état brut, mais les convictions qui prennent forme dans l'esprit de Fourniret quand il est confronté à des preuves incriminantes, des convictions du genre « On a trouvé des vêtements et on va pouvoir prouver qu'ils appartiennent à mes victimes ». On notera que le mot *preuves* est utilisé dans le passage cité ; ce mot-là renvoie sans le moindre doute aux vêtements.

Un deuxième cas est tout à fait analogue :

(34) Angélique, de son côté, a continué à tout nier en bloc, malgré les évidences, malgré les draps et les mouchoirs ensanglantés retrouvés à son propre domicile. (D, 23 novembre 2004)

Traduit en anglais, cela pourrait donner : "*Angélique, on her part, has continued to deny any wrongdoing, in spite of the evidence, in spite of the bloodied bed linen and handkerchiefs retrieved at her own home.*" Les draps et les mouchoirs constituent-ils les « évidences » dont il est question ? Le lecteur de langue anglaise risquerait de nouveau de se laisser prendre au piège. En réalité, les évidences sont les vérités qui ont pu être établies au-delà du doute raisonnable¹⁷ : à savoir qu'Angélique était avec la victime au moment des faits

17 Notre choix de mots est délibéré : sur la foi d'un dictionnaire bilingue qui traduit des termes tels que *reasonable doubt*, *reasonable chance*, *reasonable force* et *reasonable grounds* sans recourir au mot *raisonnable*, Wierzbicka (2006d : 138) fait valoir que *reasonable* et *raisonnable* sont de flagrants faux amis – ce que confirment d'ailleurs les dictionnaires spécialisés (par exemple Van Roey *et al.* 1995 : 576 ; voir aussi Geoffroy 2001 : 104-105). Il y a pourtant des occurrences de termes tels que *doute raisonnable* et *force raisonnable*, pour ne retenir que deux exemples, dans des publications aussi respectées que *Le Monde* et *Le Monde Diplomatique* : autre exemple du décalage inévitable des ouvrages de référence et des dangers qui existent quand on s'y appuie d'une façon trop hâtive. À l'appui de sa thèse, Wierzbicka fait valoir que *demande raisonnable* et *prix raisonnable* n'ont pas le même sens que *reasonable request* et *reasonable price*. Il faudrait le prouver... au-delà du doute raisonnable, ce que Wierzbicka, à notre avis, n'a pas fait. De façon plus générale, il vaudrait la peine de soumettre les adjectifs *reasonable* et *raisonnable* à un examen approfondi reposant entre autres sur des sources

et que sa version des faits ne résiste pas à l'examen. Pour ne pas parler de son comportement devant le juge ni de ses propos qui frôlent l'incohérence.

Mis à part les vêtements, les draps et les mouchoirs, qu'est-ce qui peut servir d'*evidence* ? La liste risque d'être infinie : en fait, tout ce qui est empiriquement observable et tout ce qui permet de formuler des savoirs fondés sur l'observation (plutôt que sur la raison pure et simple, à la façon des évidences françaises, qui ne sont pas empiriques) peut prétendre au statut d'*evidence*, que ce soit *evidence of* ou *in favour of* (ou *against*), *evidence for*, *evidence that*, etc. Le mot *evidence* est souvent accompagné de l'adjectif *empirical* – v. p. ex. Peeters 2006, et aussi (39) ci-dessous –, même si cet adjectif est en quelque sorte redondant, puisque toute forme d'*evidence* est en fait empirique¹⁸.

- (35) X-rays reveal clear evidence of alteration in the underdrawing of both paintings. (COBUILD)
- (36) Now, it is remarkable how generally it can be shown, by the presence of upraised organic remains, that the fringed islands have been elevated: so far, this is indirect evidence in favour of our theory. (Darwin 1845)
- (37) Many modern philosophers claim that probability is a relation between a hypothesis and the evidence for it. (Hacking 1975)
- (38) This was not the case for Ms. Schiavo, who was sufficiently evaluated by neurologists for the Florida Supreme Court to rule that there was clear and convincing evidence that she was in a persistent or permanent vegetative state. (Didion 2005)
- (39) This idealization must be kept in mind when one is considering the problem of confirmation of grammars on the basis of empirical evidence. (Chomsky 1972)

Le mot anglais *evidence* se traduit le plus souvent dans d'autres langues à l'aide d'un terme qui correspond à l'anglais *proof* ou *proofs* (Wierzbicka 2002 : 96). Mais, comme le terme anglais qu'ils traduisent d'une façon plus immédiate, ces termes-là ont un sens plus large : ils s'utilisent non seulement dans le cas de savoirs fondés sur l'observation (comme *evidence*), mais aussi dans le cas de savoirs fondés sur la raison (dans des domaines aussi divers que les mathématiques, la logique, la philosophie, etc.). Une autre différence entre *evidence* et *proof(s)* est que *evidence* est contestable, alors que des *proofs*, du moins en principe, ne se prêtent pas à la réfutation – ainsi que le montre le contraste suivant :

- (40) They plan to challenge the reliability of the evidence in front of the jury. (COBUILD)
- (41) ??They plan to challenge the reliability of the proofs in a laboratory setting.

On peut proposer les explicitations suivantes :

[F] P is evidence for / of / in favour of X =

je sais quelque chose de P

comme les gens peuvent savoir quelque chose d'autre chose

quand ils voient cette autre chose

non quand ils ne l'ont pas vue

historiques : une telle étude montrerait probablement qu'ils diffèrent désormais de façon bien moins significative qu'autrefois.

18 Les exemples (36) à (39) sont cités d'après les manuscrits de Wierzbicka auxquels font allusion les notes 8 et 11.

à cause de cela, je dis que X est vrai
 je le dis parce que je veux que vous pensiez la même chose que moi
 d'autres gens diront peut-être que X n'est pas vrai
 parce que ces gens savent quelque chose d'autre chose
 non de la même chose que les gens qui disent : X est vrai.

[G] P is evidence against X =

je sais quelque chose de P
 comme les gens peuvent savoir quelque chose d'autre chose
 quand ils voient cette autre chose
 non quand ils ne l'ont pas vue
 à cause de cela, je dis que X n'est pas vrai
 je le dis parce que je veux que vous pensiez la même chose que moi
 d'autres gens diront peut-être que X est vrai
 parce que ces gens savent quelque chose d'autre chose
 non de la même chose que les gens qui disent : X n'est pas vrai.

Une comparaison de ces deux formules avec la formule [E], répétée ci-dessous, révèle des différences aussi bien que des points communs.

[E] P, c'est une évidence =

je dis P
 je sais que P est vrai
 je pense :
 les gens ne peuvent pas ne pas savoir que P est vrai
 à cause de cela, quand les gens pensent à P, ils ne peuvent pas dire que
 P n'est pas vrai.

C'est sur les différences qu'on insiste le plus souvent quand on rapproche deux faux amis – et ces différences sont bien entendu clairement marquées dans les explicitations ci-dessus. *Evidence*, en anglais, est avancé en vue de convaincre quelqu'un du bien-fondé (ou du « mal-fondé ») d'une thèse (le « X » dans [F] et [G]), alors que dans le cas d'une *évidence* à la française il n'y a plus rien à prouver. *Evidence*, en anglais, est un acquis fondé sur l'observation (ou sur un autre processus inférentiel), alors que dans le cas d'une *évidence* à la française on a affaire à un donné. *Evidence*, en anglais, peut être contesté, alors qu'une *évidence* à la française est perçue comme indéniable. Cependant, à côté des différences, il y a aussi des points communs qu'il convient – qu'on nous pardonne le jeu de mots – de mettre en évidence. On n'en signalera ici que les plus importants, que le recours à la MSN fait d'ailleurs sauter aux yeux. Dans les deux cas, un savoir est en cause (« je sais quelque chose de P » *versus* « je sais que P est vrai »). Dans les deux cas, un pensé est en cause, que ce soit un pensé que l'on cherche à provoquer (« je veux que tu penses la même chose que moi ») ou un pensé déjà établi (« je pense : les gens ne peuvent pas... »). Enfin, dans les deux cas, il est question de vérités, acceptées ou réfutées (« je dis que X est vrai », « je dis que X n'est pas vrai ») ou bien intuitives (« je sais que P est

vrai »). Si on insiste bien moins souvent sur les points communs que sur les différences, c'est que d'ordinaire ils échappent à l'observation, dans la mesure où ce sont... des évidences. Mais comme toutes les évidences, celles-là aussi méritent d'être rappelées de temps en temps, ne fût-ce que pour corroborer la thèse que la relation précise entre deux faux amis est le plus souvent – et de toute façon nettement plus souvent qu'on ne l'a cru jusqu'ici – une question de degré.

Concluons. Que ce soit dans le domaine de la sémantique, de la syntaxe, de la pragmatique, de la lexicologie ou de la lexicographie, l'approche MSN a d'ores et déjà fait ses preuves. L'objectif n'a jamais varié : il s'agit d'aller au-delà des descriptions traditionnelles, où abondent les termes techniques, obscurs, ambigus, complexes et culturellement spécifiques (c'est-à-dire ethnocentriques), en formulant des métacommentaires plus transparents qu'en français qu'on pourrait appeler des *explicitations sémantiques* (d'après l'anglais *semantic explications*). On parle aussi de *formules sémantiques* – et, dans l'étude de l'interculturel, de *scénarios culturels* (angl. *cultural scripts* ; le terme français a été proposé par Hagège 1998). Les explicitations et les scénarios prennent la forme de séquences de propositions exprimées en MSN. Détaillés, rigoureux et exhaustifs, ils permettent de décrire les faits de langue les plus divers (et aussi d'élucider les termes techniques utilisés pour en rendre compte). Grâce à l'universalité des éléments constitutifs de la métalangue, les explicitations et les scénarios sont traduisibles, tels quels et sans la moindre ambiguïté, dans toutes les langues du monde¹⁹. Contrairement aux descriptions souvent obscures de ceux qui ne se méfient pas assez de la spécificité de leur jargon, ils sont vérifiables par des locuteurs natifs, à qui on peut sans aucune difficulté, sans courir le risque qu'ils s'y trompent, soumettre des paraphrases rédigées en MSN. Instrument efficace dans la lutte contre le « syndrome du franchissement du gué »²⁰, la MSN est indispensable à tous ceux qui ont souci de bien se faire comprendre.

Notre but, en recourant à la MSN, a été d'encourager une approche plus subtile des faux amis que celle dont on a l'habitude. Certes, il y a des cas où une perspective plus radicale reste entièrement légitime. Ils sont de deux types. Le premier se laisse illustrer à l'aide des mots *seize* « 16 » en français et *seize* « saisir » en anglais (v. plus haut) ; ce sont des mots qui n'appartiennent même pas à la même partie du discours²¹. Le second est plus surnois ; pensons à *déterrer* en français et *deter* en anglais. En dépit du fait qu'il s'agit de verbes, ils n'ont absolument rien à voir l'un avec l'autre. *Déterrer*, on le sait, veut dire

19 À moins que, pour ne pas trop alourdir une explicitation, elles comportent des *molécules sémantiques*, c'est-à-dire des termes sémantiquement complexes qui ne sont pas nécessairement universels et qu'il faut dès lors garder intacts lors de la transposition (par exemple la molécule *nez_M*, dans une explicitation du mot *mouchoir* ; ces molécules sont identifiées à l'aide du M souscrit qui les suit). Les molécules ne sont admises que dans des conditions très précises, et doivent être explicitées séparément. Pour plus d'informations, on verra Peeters (2010).

20 « Ce syndrome, relevé pour la première fois au Moyen-Âge chez les moines de l'Oelenberg (à Reiningue, près de Mulhouse), est bien connu : on continue de sauter d'une pierre à une autre, sans jamais tomber à l'eau, mais on oublie de franchir la rivière ! » (Kleiber 2001 : 3).

21 On peut se poser la question de savoir s'il s'agit vraiment de faux amis. Walter (2001) semble croire que oui. Elle avait déjà proposé plusieurs autres exemples du même type dans Walter (1994 : 422-423), ironiquement dans une rubrique intitulée « Méfions-nous des apparences ».

« faire sortir de la terre » ; le verbe anglais correspondant, formé d'après les mêmes principes et de façon tout aussi transparente, est *unearth*. *Deter*, d'autre part, n'a rien à faire avec la terre, mais doit être rapproché de mots tels que *terror* et *terrorise*. Il signifie « décourager, dissuader (par exemple en menaçant de semer la terreur) » (v. annexe 2).

Des exemples tels que *seize* / *seize* (qui ne sont peut-être pas à retenir ; v. note 21) et *déterrer* / *deter* constituent les exceptions qui confirment la règle. Il y a lieu de croire que derrière la grande majorité des faux amis se cachent des similarités sémantiques qui risquent d'échapper à la vigilance du lexicographe, à moins qu'il ne se serve d'un outil descriptif aussi précis que la MSN. Celle-ci s'utilise le plus souvent pour expliciter des différences que le recours à des termes trop complexes tend à masquer. Or, nous espérons avoir montré qu'elle se prête également à la recherche de similarités, dont la reconnaissance permet de mieux comprendre la nature du rapport qui existe entre des faux amis particuliers. Ceux-ci ne sont pas toujours « aussi faux », ni même toujours « aussi partiels », qu'il y paraît. C'est une constatation qui a sans doute des conséquences pédagogiques qu'il vaudrait la peine d'examiner. Dans le cadre de cette présentation, il a été impossible de développer cet aspect.

ANNEXE 1 – CONTEXTES FRANÇAIS OÙ FIGURE LE MOT *ÉVIDENCE* DANS LE CORPUS DÉPOUILLÉ (nombre d'occurrences entre parenthèses)

mettre en évidence (13)
à l'évidence (12)
de toute évidence, d'évidence (7)
devant l'évidence / les évidences (4)
c'est une évidence (2)
apposer X en évidence (1)
éprouver des évidences (1)
les évidences font défaut (1)
malgré les évidences (1)
se rendre à l'évidence (1)
il reste une évidence (1)
tenir de l'évidence (1)
X en évidence [construction appositive] (1)

ANNEXE 2 – DESCRIPTION (TRÈS APPROXIMATIVE) EN MSN DES VERBES *DÉTERRER* (FRANÇAIS) ET *DETER* (ANGLAIS)

X *déterre* Y =

X fait quelque chose

à cause de cela, les gens peuvent voir toutes les parties de Y

avant cela, les gens ne pouvaient pas voir toutes les parties de Y

parce que beaucoup de parties de Y étaient dans la terre_M

X *deters* Y =

Y veut faire quelque chose

X ne veut pas que Y fasse cela

X fait autre chose

à cause de cela, Y ne peut plus faire ce qu'il voulait faire

Les deux éléments communs (« faire quelque chose » et « à cause de cela ») ne sont pas de nature à compromettre la thèse d'une différence irréductible : le premier concerne en fait deux « entités » différentes (le sujet X et l'objet direct Y) et, dans le cas du verbe anglais, est précédé du verbe modal *vouloir* qui transforme le « faire » en un « vouloir faire » ; le deuxième est trop abstrait pour justifier, à lui seul, l'idée que les deux mots auraient malgré tout des sens rapprochables.

RÉFÉRENCES BIBLIOGRAPHIQUES

- CARROLL Raymonde, 1987, *Évidences invisibles : Américains et Français au quotidien*, Paris, Le Seuil.
- CARROLL Raymonde, 1988, *Cultural Misunderstandings: The French-American Experience*, Chicago, The University of Chicago Press.
- GEOFFROY Christine, 2001, *La Mésentente cordiale : Voyage au cœur de l'espace interculturel franco-anglais*, Paris, Grasset et Le Monde.
- GODDARD Cliff (ed.), 2008, *Cross-Linguistic Semantics*, Amsterdam, Benjamins.
- GODDARD Cliff and WIERZBICKA Anna (eds.), 1994, *Semantic and Lexical Universals: Theory and Empirical Findings*, Amsterdam, Benjamins.
- GODDARD Cliff and WIERZBICKA Anna (eds.), 2002, *Meaning and Universal Grammar: Theory and Empirical Findings*, Amsterdam, Benjamins (2 vol.).
- GRANGER Sylviane et SWALLOW Helen, 1988, "False friends: A kaleidoscope of translation difficulties", *Le Langage et l'Homme*, 23, p. 108-120.
- HAGÈGE Claude, 1998, Compte rendu de M. Hellinger et U. Ammon (éd.), *Contrastive sociolinguistics* (Berlin, Mouton de Gruyter, 1996), *Bulletin de la Société de linguistique de Paris*, 93 (2), p. 102-105.
- KLEIBER Georges, 2001, « Les référents évolutifs : Au large et à l'étroit », dans W. De Mulder et C. Schnedecker (éds), *Les Référents évolutifs entre linguistique et philosophie*, Metz, Université de Metz, p. 3-37.
- PEETERS Bert, 1994, "Semantic and lexical universals in French", in C. Goddard et A. Wierzbicka (eds.), *Semantic and Lexical Universals: Theory and Empirical Findings*, Amsterdam, Benjamins, p. 423-442.
- PEETERS Bert, 1997, "The syntax of time and space primitives in French", *Language sciences*, 19, p. 235-244.
- PEETERS Bert (éd.), 2006, *Semantic Primes and Universal Grammar: Empirical Evidence from the Romance Languages*, Amsterdam, Benjamins.
- PEETERS Bert, 2007, « Le stéréotypage du tueur en série dans les faits divers de la presse écrite : Le cas de Michel Fourniret », dans H. Boyer (éd.), *Stéréotypage, stéréotypes : Fonctionnements ordinaires et mises en scène, tome 1, Média(tisation)s*, Paris, L'Harmattan, p. 241-251.

- PEETERS Bert, 2009, « Tu ou vous ? », dans B. Peeters et N. Ramière (éds), *Tu ou vous : L'embarras du choix*, Limoges, Lambert-Lucas, p. 9-39.
- PEETERS Bert, 2010, « La métalangue sémantique naturelle : Acquis et défis », dans J. François (éd.), *Grandes Voies et chemins de traverse de la sémantique cognitive, Mémoires de la Société de linguistique de Paris*, N.S., 18, p. 75-101.
- VAN ROEY Jacques, GRANGER Sylviane et SWALLOW Helen, 1995, *Dictionnaire des faux amis français / anglais*, Louvain-la-Neuve, Duculot.
- WALTER Henriette, 1994, *L'Aventure des langues en Occident : Leur origine, leur histoire, leur géographie*, Paris, Laffont.
- WALTER Henriette, 2001, « Les “faux amis” anglais et l'autre côté du miroir », *La Linguistique*, 37 (2), p. 101-112.
- WIERZBICKA Anna, 1972, *Semantic Primitives*, Frankfurt, Athenäum.
- WIERZBICKA Anna, 2002, “Philosophy and discourse: The rise of ‘really’ and the fall of ‘truly’”, *Cahiers de praxématique*, 38, p. 85-112.
- WIERZBICKA Anna, 2006a, « Les universaux empiriques du langage : tremplin pour l'étude d'autres universaux humains et outil dans l'exploration de différences transculturelles », *Linx*, 54, p. 151-179.
- WIERZBICKA Anna, 2006b, « Sens et grammaire universelle : théorie et constats empiriques », *Linx*, 54, p. 181-207.
- WIERZBICKA Anna, 2006c, “Anglo scripts against ‘putting pressure’ on other people and their linguistic manifestations”, in C. Goddard (ed.), *Ethnopragmatics: Understanding Discourse in Cultural Context*, Berlin, Mouton de Gruyter, p. 31-63.
- WIERZBICKA Anna, 2006d, *English: Meaning and culture*, Oxford, Oxford University Press.
- WIERZBICKA Anna, 2010, *Evidence, Experience, Sense: The Hidden Cultural Legacy of English*, Oxford, Oxford University Press.

LA TRADUCTION EN LEXICOGRAPHIE BILINGUE : DES MARQUES DE CATÉGORISATION IMPLICITES ?

par Christophe CUSIMANO
Université de Metz - Ceted

Résumé — Quand ils écrivent des dictionnaires bilingues de langues de statut sociolinguistique inégal, les lexicographes dont la langue est la plus prestigieuse sont influencés par leurs propres habitudes culturelles dans leurs choix de traductions. Notre étude présente une linguiste allemande qui a étudié une langue amérindienne, le kaingang, dont elle a élaboré un dictionnaire, et compare les traductions de ce dictionnaire avec nos propres résultats et données de terrain.

Mots clés — kaingang, dictionnaire bilingue, transfert culturel, catégories and étiquettes sémantiques.

Abstract — When writing bilingual dictionaries that concern languages of different sociolinguistic status, the lexicographers belonging to the more prestigious language are influenced by their own cultural habits in their translation choices. This study looks at a German linguist who described an Amerindian language, the Kaingang, for which she wrote a dictionary, and contrasts the translations given by the dictionary with our own results and fieldwork data.

Keywords — kaingang, bilingual dictionary, cultural transfer, semantic categories and labels.

INTRODUCTION

Comme nous le savons, le but de tout dictionnaire bilingue est de donner de deux langues des traductions terme à terme les plus fidèles possibles. Or cette démarche de traduction, qui s'accompagne d'une lourde perte en termes de nuances sémantiques, tend souvent à une série d'analogies qui révèlent les références culturelles du lexicographe. Ainsi le dictionnaire bilingue livre-t-il au lecteur, bien à son insu, des marques de catégorisation implicites. C'est ce que nous nous proposons de démontrer à propos du *Dictionnaire bilingue kaingang / portugais* d'U. Wiesemann (2002), missionnaire allemande, par la traduction en portugais de certains composés sémantiques (Creissels 2004 : 23) du kaingang qui définissent un sentiment.

1. PRÉSENTATION DU KAINGANG

L'extrait ci-dessous de l'*Enciclopedia das línguas no Brasil* replace le kaingang au sein des langues du monde et des langues amérindiennes :

Langue indigène de la famille Jê, qui fait partie de la branche Macro-Jê (cf. Rodrigues 1986). Parlé au Sud du Brésil, le kaingang forme avec le xokleng l'ensemble méridional restreint des langues Macro-Jê : ce sont les dernières d'une quinzaine de langues Jê, qui s'étendent au sud du tropique du Capricorne. D'un autre côté, le kaingang est la langue Jê qui a le plus grand nombre de locuteurs, la population kaingang correspondant à plus de 45 % de la population totale des peuples de langue Jê.

Les Kaingangs habitent une trentaine de réserves situées dans une zone comprise entre le fleuve Tietê (limite septentrionale de l'état de São Paulo) et le fleuve Ijuí (limite méridionale du Rio Grande do Sul). La langue kaingang présente une variété de dialectes. Wiesemann (1971 : 259-260) en identifie cinq : (a) São Paulo ; (b) Parana, entre les fleuves Paranapanema e Iguaçu ; (c) Central, entre les fleuves Iguaçu et Uruguay ; (d) Sud-est, au sud du fleuve Uruguay et à l'ouest du fleuve Passo Fundo (plus quelques familles à Votouro) ; (e) Sud-est, au sud du fleuve Uruguai et à l'est du fleuve Passo Fundo. De fait, la richesse dialectale kaingang est encore plus grande, en partie en raison des migrations et des alliances historiques entre les villages. (traduit de la page <http://www.labeurb.unicamp.br/elb/indigenas/kaingang.htm>, Université de Campinas)

Le kaingang est donc une langue amérindienne parlée dans le sud du Brésil et c'est dans la réserve Cambuí à São José dos Pinhais, proche de Curitiba (capitale de l'État du Paraná) que nous avons mené nos recherches. Il faut préciser que le kaingang est une langue ergative qui se comporte en outre selon tous les critères d'une langue isolante : les unités lexicales sont formellement immuables et aucune flexion ne les affecte. Ce caractère isolant, qui marque selon W. von Humboldt (1859) l'appartenance d'une langue à la phase II de sa théorie de l'agglutination – la première n'étant qu'un stade hypothétiquement réservé à la désignation d'unités concrètes –, aura une grande importance dans notre raisonnement. Enfin, il convient d'ajouter que le kaingang est ce que l'on a coutume d'appeler une langue à tradition orale : les règles orthographiques ont été fixées provisoirement par un groupe de travail mixte portugais - kaingang en 1976, avant d'être confirmées en 1997. Par conséquent, il n'est pas étonnant que mis à part le Nouveau Testament, nous ne disposions pas dans cette langue de textes susceptibles d'étayer notre raisonnement, mais seulement d'énoncés. On ne dispose que de peu de sources écrites, d'où l'intérêt d'enquêtes de terrain.

2. LA « COMPOSITION SÉMANTIQUE »

Comme nous l'avons laissé entendre, cette langue offre une illustration intéressante du phénomène universel de composition, et plus particulièrement de composition sémantique que Creissels définit de la sorte :

On peut parler de composition sémantique lorsqu'une combinaison de mots pleins qui a l'apparence d'un syntagme formé selon les règles productives de construction de syntagmes est régulièrement utilisée avec un signifié qu'on ne peut pas prédire en appliquant simplement les règles sémantiques qui en principe permettent de calculer le signifié d'un syntagme à partir du signifié des mots qui le constituent. (Creissels 2004 : 23)

Ces composés seraient donc plutôt le résultat d'habitudes langagières bien ancrées, comme le sont en français « cœur de pierre », « cœur d'ange », ou encore « sourire d'ange ». Ils prennent appui sur ce que nous avons appelé des morphèmes génériques¹, dont le signifié est d'importance majeure dans la langue en question, qui servent de base sémantique pour la composition et en déterminent le genre, comme *fe* dont nous allons traiter à présent.

3. « ÉLANS DU CŒUR » OU « ÉTATS D'ESPRIT » ?

En kaingang, de nombreux termes dont le signifié définit un sentiment contiennent le segment phonique /fe/, *fe* selon les règles orthographiques de la langue, la plupart du temps à l'initiale. Le morphème *fe* se traduit par « sein, poitrine ou cœur » (1a) et (1b) (énoncé d'une locutrice kaingang, Wiesemann 2002 : 13). Il peut aussi signifier, selon nos informateurs, « centre des sentiments » ou même « centre universel des sentiments ».

(1a) Porko fe ko mē sóg nĩ
 porc poitrine manger beaucoup moi avoir
 « J'ai mangé beaucoup de poitrine de porc »

(1b) Kejẽn fe kanhró ěg nĩ
 Parfois cœur savoir nous avoir
 « Parfois nous savons par le cœur »

Ce polysème fonctionne comme base sémantique du composé, comme dans *fe mrir*, littéralement « cœur ressuscité » :

(1c) Fe mrir inh nĩ
 ? moi, je avoir

Nous allons en outre nous appuyer sur deux autres exemples, *fe pẽ* et *krĩ ku-prã*, dont les traductions sont à notre avis représentatives des dérives lexicographiques dont il est ici question.

Fe mrir tout d'abord est traduit « *vivo, alegre* » par U. Wiesemann, c'est-à-dire « vif, joyeux » (pour distinguer son texte du nôtre, nous reproduisons ses exemples en facsimilé) :

fe mrir *sub.dep.* vivo, alegre. **Fe mrir ti nĩ ha. Kĩ tóg**
jũ mãn tũ nĩ, ã tỹ kron mũ ti. Aquele que bebia,
 agora não fica mais brabo mas vive com alegria.

C'est déjà en soi une sur-interprétation des données linguistiques, puisque ce n'est pas la traduction qu'en donnent les locuteurs de kaingang, qui lui préfèrent celle, littérale, de « cœur ressuscité ». Toutefois, cette traduction a le mérite de laisser entendre que ce composé exprime en kaingang un *sentiment*, comme c'est aussi le cas pour *fe pẽ* (2), traduit par « *amor* » (« amour » en français) :

fe pẽ *sub.dep.* amor para com o filho. **Inh fe pẽ vẽ,**
inh kósĩn ti. Tenho muito amor para com o meu
 filho.

¹ Ils ne correspondent pas à ce que l'on regroupe communément sous l'étiquette « classificateurs » car, pour reprendre la terminologie de Creissels, ils effectuent bien des *opérations sémantiques* sur les termes qu'ils accompagnent.

- (1d) Fe pẽ ti nĩ
 cœur beaucoup, vrai elle, son avoir
 « Elle a beaucoup d'amour »

Comme nous l'avons suggéré, il semble en effet possible de considérer les composés construits depuis *fe* comme des sentiments, des *élans du cœur* en somme. Mais comme lorsqu'en (2), apparaît un autre composé, *krĩ kuprã*,

- (2) Krĩ kuprã inh nĩ ha, Topẽ tugrĩn
 tête vide je avoir maintenant Dieu à cause de, grâce à
 « Je suis en paix maintenant, grâce à Dieu »

formé à partir de *krĩ* «tête», la traduction «*em paz, contente*» («en paix, content») laisse de nouveau entendre qu'il s'agit d'un terme de sentiments. Mais sa traduction signifie clairement que le lexicographe n'a pas compris que les composés dont la base est *krĩ* sont plutôt des états d'esprit, ce qui constitue une différence remarquable.

krĩ kuprã *sub.dep.* em paz, contente. **Krĩ kuprã inh**
nĩ ha, Topẽ tugrĩn. Tenho paz por causa de Deus.

Wiesemann commet la même confusion à propos de *krĩ kusa*, lit. «tête froide», qu'elle choisit de traduire, à tort nous semble-t-il, par «en paix, sans souci».

krĩ kusa *sub.dep.* em paz, despreocupado. **Krĩ kusa**
inh nĩ ha. Agora estou em paz, despreocupado.

CONCLUSION

Voici donc brièvement illustré comment, selon nous, un dictionnaire bilingue peut neutraliser, par ses traductions, les propriétés catégorielles d'une des deux langues, de culture non dominante et éloignée (selon un critère d'éloignement culturel bien sûr tout relatif), au profit des marques de catégorisation en vigueur dans la langue du lexicographe. Parmi les conséquences majeures, la catégorisation ainsi induite, ne rendant nullement compte des particularités sémantiques de la langue traduite, ne permet pas de reconstituer un réseau lexical fidèle.

Il serait de loin préférable de conserver dans la définition la mention du sens littéral des éléments du composé : le travail de traduction à effectuer en moins, source potentielle d'ethnocentrisme sémantique, nous proposons d'effectuer un marquage lexicographique plus précis, comportant notamment des informations morphologiques (relatives au figement des compositions et expressions), et taxémiques, ou pour le dire plus simplement, de champ sémantique. Une telle entrée de dictionnaire kaingang / français pourrait ressembler à ceci :

fe pẽ comp. sém. («cœur, centre des sentiments» (base / M. G.) «beaucoup, vrai») – *ch. sém.* : sent. – Souvent employé dans le sens d'«amour». *Fe pẽ ti nĩ*, «Elle a un cœur vrai / Elle a de l'amour».

comp. sém.	composé sémantique
base	base du composé
M. G.	morphème générique
ch. sém.	champ sémantique
sent.	sentiments.

À défaut d'être parfaitement adéquate, une telle présentation permettrait d'éviter les confusions entrevues dans cet article.

RÉFÉRENCES BIBLIOGRAPHIQUES

- CUSIMANO Christophe, 2006, « Os substantivos de sentimentos em kaingang », en ligne : <http://www.sil.org/americas/brasil/PUBLICNS/LING/KGSubSnt.pdf>.
- CREISSELS Denis, 2004, *Cours de syntaxe 2004, Chapitre 1, Les notions de base de l'analyse syntaxique*, <http://lesla.univ-lyon2.fr/IMG/pdf/doc-346.pdf>.
- CREISSELS Denis, 2004, *Cours de syntaxe 2004, Chapitre 2, Genres, classes nominales, classificateurs*, <http://lesla.univ-lyon2.fr/IMG/pdf/doc-368.pdf>.
- HUMBOLDT Wilhelm Von, 1859, *De l'origine des formes grammaticales et de leur influence sur le développement des idées, suivi de Lettre à M. Abel Rémusat*, Éditions Ducros.
- WIESEMANN Ursula Gojtéj, 2002, *Dicionário bilingüe kaingang / português*, Curitiba, Editora Evangélica Esperança.
- Enciclopedia das línguas no Brasil :
<http://www.labeurb.unicamp.br/elb/indigenas/kaingang.htm>

LES POINTS DE VUE DANS LE LEXIQUE ET DANS LE DICTIONNAIRE

par Haïfa BEN MAHFOUDH-HUBERT, Erzsébet CHMELIK
et Pierre-Yves RACCAH (CNRS)
CeReS, Université de Limoges

Résumé — Ce travail s’inscrit dans un programme de recherche dont le but est de fournir un cadre théorique efficace et fiable pour la construction d’un dictionnaire de langue qui tienne compte des *présupposés idéologiques cristallisés dans le lexique*, c’est-à-dire des points de vue dont les mots de la langue imposent l’acceptation préalable à la compréhension des discours et des textes qui les utilisent. L’objectif de l’article est de montrer qu’il est possible de fournir, de manière systématique, des éléments de description des points de vue présupposés dans les mots du lexique : la première partie montre que les lexicographes jugent nécessaire d’inclure, dans un article de dictionnaire, des indications sur la *manière* dont les mots désignent, mais fournissent ces indications de manière non systématique ; la deuxième partie montre que cette *manière* dont les mots désignent s’appuie sur des *présupposés idéologiques* cristallisés dans les mots et pouvant être décrits en termes de contraintes sur les *points de vue* que l’usage de ces mots oblige à adopter ; à l’occasion d’un exemple, la troisième partie illustre la méthode et justifie les choix.

Mots clés — description sémantique, description lexicale, points de vue lexicalisés, présupposés idéologiques, argumentation, lexicographie.

Abstract — This article is a contribution to a research programme whose aim is to provide an efficient and reliable theoretical framework for dictionaries of natural languages which account for the *ideological presupposition crystallised within the lexicon*: these dictionaries are expected to describe the *points of view* imposed by the words as previous agreements in order to understand the texts and discourses which use those words. The intention of the authors is to show that it is possible to provide, in a systematic way, the essential elements for the description of the points of view presupposed in the words of the lexicon. The first section shows that lexicographers, themselves, find it necessary to give indications on *how* words refer to what they refer to; however, it will be shown that these indications are given in a non systematic way. In section 2, we show that *how words refer* relies on ideological presuppositions which they crystallise, and which can be accounted for in terms of constraints on the *points of view*

committed with the use of those words. In section 3, while presenting an example, we illustrate and discuss that method, and give justifications for the choices that have been made.

Keywords — semantic description, lexical description, lexicalised points of view, ideological presupposition, argumentation.

INTRODUCTION

Le présent travail s'inscrit dans un programme de recherche dont le but est de fournir un cadre théorique efficace et fiable pour la construction d'un dictionnaire de langue qui tienne compte des *présupposés idéologiques cristallisés dans le lexique*, c'est-à-dire des points de vue dont les mots de la langue imposent l'acceptation préalable à la compréhension des discours et des textes qui les utilisent. L'objectif de cet article est très modeste au regard du programme dans lequel il s'inscrit : nous voulons montrer qu'il est possible de fournir, de manière systématique, des éléments de description des points de vue présupposés dans les mots du lexique, et ce, en utilisant un modèle descriptif de ce que l'on peut appeler une « sémantique instructionnelle ». Nous n'avons pas l'ambition de montrer ici que ce modèle est meilleur qu'un autre ; en revanche, ceux que nous aurons convaincus attendront, des autres modèles, qu'ils permettent aussi des descriptions systématiques de ces phénomènes de sémantique lexicale.

Dans la première partie de l'article, après avoir qualifié les tâches relevant, respectivement, de la lexicographie, de la lexicologie et de la sémantique lexicale (tâches qui peuvent être – et qui sont souvent – cumulées par une même personne), nous montrons que le noyau *systématique* des descriptions fournies par les dictionnaires concerne, actuellement, la dénotation. Nous verrons en quoi et pourquoi ces indications dénotationnelles relatives aux mots ne sont pas suffisantes : les lexicographes eux-mêmes jugent nécessaire d'inclure, dans un article de dictionnaire, des indications sur la *manière* dont les mots désignent. Ainsi, les principaux dictionnaires donnent souvent une partie de ces indications, mais de manière non systématique, au moyen d'exemples.

Dans la deuxième partie de l'article, nous montrons que cette manière dont les mots désignent s'appuie sur des *présupposés idéologiques cristallisés* dans les mots et propres à chaque langue. Cette métaphore de la cristallisation vise un phénomène bien connu : pour comprendre un discours, il est souvent (toujours ?) nécessaire d'admettre, au préalable, certains points de vue sur ce dont le discours parle, mais aussi sur d'autres entités ; ces points de vue implicites et nécessaires à la compréhension sont imposés par le choix des mots de la langue que le discours met en œuvre. Nous proposons un ensemble de notions permettant de décrire ces présupposés idéologiques cristallisés. Après avoir précisé ce que l'on doit attendre d'une description sémantique, et justifié ce « cahier des charges », nous montrons que la signification des mots peut être décrite en termes de contraintes sur les *points de vue* que l'usage des mots oblige à prendre en compte.

Dans la troisième partie, nous illustrons l'ensemble des réflexions présentées au moyen d'un exemple détaillé, destiné à la fois à compléter la justification du cahier des charges et à montrer comment les options choisies s'en acquittent.

1. LEXICOLOGIE, LEXICOGRAPHIE ET SÉMANTIQUE : QUELQUES DIFFICULTÉS

Nous commençons par caractériser le rôle d'un dictionnaire de langue, puis nous examinons la question de savoir si, et dans quelle mesure, les moyens utilisés actuellement par les dictionnaires pour satisfaire ce rôle sont adéquats. À l'occasion de l'examen de ce rôle, on signalera, sans s'y attarder, les notions qui, bien que régulièrement utilisées par tout sujet parlant – linguiste ou non-linguiste – sont trop vagues pour que leur utilisation nous permette une analyse rigoureuse : elles feront l'objet de caractérisations plus précises, dans la deuxième section.

1.1 À PROPOS DU RÔLE D'UN DICTIONNAIRE

La principale attente d'un utilisateur de dictionnaire est de trouver, dans un tel ouvrage, une description de la signification d'un mot, description qui soit telle qu'il puisse utiliser ce mot de façon appropriée, c'est-à-dire comprendre un discours utilisant une réalisation de ce mot, ou produire un discours dans lequel une occurrence de ce mot intervient¹. Pour que ce rôle des dictionnaires soit précisé, il faut, bien sûr, que l'on comprenne de manière claire en quoi la *signification* d'un mot intervient dans l'interprétation d'un discours utilisant ce mot (que ce soit en production ou en compréhension)². C'est ce que nous verrons au début de la deuxième section.

Dès à présent, cette caractérisation fait apparaître qu'un dictionnaire se situe entre *discours* et *langue*³. Ainsi, les auteurs de dictionnaires sont amenés à travailler dans cette interface : ils cherchent à fournir des descriptions d'unités de langue permettant d'utiliser ces unités dans la construction et dans la compréhension des discours. Pour le lexicologue, l'objectif est

d'établir la liste des unités qui constituent le lexique, et de décrire les relations entre ces unités. (Martin-Berthet & Lehmann 2000 : 13)

en tenant compte de ce que

la lexicologie et la lexicographie se sont constamment inspirées mutuellement, mais l'objet « langue » et l'objet « dictionnaire » sont de nature différente. (*ibid.* : 14)

Ainsi, on peut voir la lexicologie comme « l'étude analytique des faits de vocabulaire » (Matoré 1953 : 13), ou encore l'étude « non seulement [de] tous les mots attestés d'une langue, mais aussi [de] tous les mots potentiellement "attestables" » : l'accent est mis sur l'observation des discours afin d'en tirer des caractéristiques des mots de la langue⁴.

1 Voir, par exemple, Gaudin & Guespin (2000 : 99) : « ... [le dictionnaire] est un ouvrage que l'on consulte principalement pour découvrir des mots inconnus, des sens nouveaux, mais aussi pour connaître le bon usage, les orthographes exactes et les significations précises des mots. Il conjugue une fonction de description et une fonction de prescription. »

2 « [dans les dictionnaires de langue,] les mots sont étudiés pour eux-mêmes ; on y étudie le fonctionnement du mot en discours, les significations qui lui sont attachées, les emplois qu'il permet, les tournures dans lesquelles il entre, etc. » (*ibid.* : 100).

3 V. Mortureux (1997 : 9) : « La première tâche de la lexicologie est de définir son objet propre ; pour cela, elle doit dégager de l'ensemble des mots observables dans les discours les unités qui constituent le matériel lexical de la langue. »

4 V. Serasset & Polguère (1997) : « le lexicologue définit les informations qui seront contenues dans le lexique, spécifie leurs formes et donne les critères permettant de définir les unités du lexique. »

Le lexicographe, dont l'objectif « consiste en la confection de dictionnaires » (Kremer 2008), « construit le lexique selon les spécifications [du lexicologue et de l'informaticien] » (Serasset & Polguère 1997). Par ailleurs, pour que les dictionnaires puissent satisfaire le rôle indiqué, les lexicographes, dans leur travail d'auteurs de dictionnaire, et, en amont, les lexicologues, dans leur travail de spécification, sont amenés à s'appuyer sur le travail des sémanticiens⁵. Pour que cette indication ne soit pas une simple pétition de principe le rôle de la sémantique devra être précisé : c'est ce que nous ferons aussi au début de la section 2.

Sans attendre ces précisions, nous pouvons déjà examiner les descriptions que les dictionnaires proposent actuellement, en insistant sur le contraste que leurs articles présentent entre une très fréquente mention des points de vue implicites évoqués par les mots et le manque de moyens de systématisation de cet aspect de la description.

1.2 À PROPOS DES MOYENS LEXICOGRAPHIQUES...

Pour faciliter l'étude de ce contraste entre le souci des lexicographes de décrire les points de vue implicites suggérés par les mots et l'absence des moyens de systématiser cette description, nous partons de l'examen d'un article spécifique dans trois dictionnaires de langue : *Larousse*, *Petit Robert*, *Robert Culturel* ; nous y relevons différents aspects sémantiques sur lesquels les lexicographes attirent l'attention de l'utilisateur des dictionnaires, et montrons des défauts de systématisation, en insistant sur le fait qu'ils ont été remarqués par les auteurs des articles, qui ont tenté d'y remédier. À partir de cet exemple, nous formulons la problématique de manière plus abstraite, laissant au lecteur le soin de vérifier que la généralisation qui s'ensuit n'est pas abusive⁶. L'entrée que nous avons choisie pour illustrer ces défauts est celle du mot *riche*⁷, qui a fait l'objet de plusieurs études approfondies⁸.

Les trois dictionnaires présentent d'abord l'origine du mot et définissent ensuite trois caractéristiques majeures que nous résumons dans le tableau 1.

Les définitions sont illustrées par des énoncés, des citations, des expressions, des proverbes... dans lesquels on trouve différents emplois du mot *riche*. Chacune de ces définitions caractérise des entités du monde : des personnes (« qui a de la fortune »), des événements ou des objets (« qui annonce ou suppose la richesse »).

1.3 LE RÔLE DES EXEMPLES

Le recours à des exemples n'a pas pour seule fonction d'illustrer les définitions : il permet souvent de les compléter : « Un *riche* capitaliste → *Gros* » (*Robert Culturel* et *Petit Robert*). D'autres cas similaires soulignent que le lexicographe

5 « Pour construire une entrée de dictionnaire, le lexicographe doit rechercher les différents usages de l'entrée donnée. Il ne peut se contenter de son "intuition" linguistique » (*ibid.*)

6 Ce manque de rigueur (généralisation à partir d'un exemple), atténué par l'évidence avec laquelle la généralisation s'impose, sera compensé par le fait que cette « simplification de la procédure » nous permettra de traiter de manière plus approfondie ce que nous proposons pour systématiser la description des points de vue implicites.

7 Nous utilisons la même notation (guillemets ordinaires) pour les mots-de-langue et pour les mots-de-discours (voir ci-dessous § 2.1), afin de faciliter la compréhension des lecteurs qui n'auraient pas intégré cette distinction dans leur appareil conceptuel effectif.

8 Voir, entre autres, une première étude dans Bruxelles *et al.* (1995) et une étude contrastive dans Raccah (1997).

a jugé utile d'indiquer une relation sémantique entre *riche* et *gros*, valable lorsque ces adjectifs qualifient le substantif *capitaliste*. De cela, on peut tirer deux conséquences, que nous développerons plus loin :

1. en premier lieu, si le lexicographe a jugé nécessaire d'évoquer cette relation sémantique, il doit la considérer comme constitutive de la signification de l'adjectif *riche* (une infinité de relations sémantiques possibles n'ont pas été évoquées, et seules quelques-unes apparaissent) ;
2. en second lieu, si le lexicographe a jugé nécessaire d'évoquer cette relation au moyen d'un exemple, c'est qu'il considère que la définition qu'il propose n'est pas suffisante pour faire apparaître cette relation (un grand nombre de citations possibles pourraient illustrer la définition, mais celles qui sont mentionnées sont, précisément, celles qui ajoutent quelque chose à ce que la définition indique).

LAROUSSE	PETIT ROBERT	ROBERT CULTUREL
1. Se dit d'une personne qui possède une grande fortune, de grands biens.	1. Qui a de la fortune, possède des biens, de l'argent en abondance.	1. Qui a de la fortune, possède des biens, de l'argent en abondance ou plus que la majorité des membres de sa communauté.
2. Se dit de ce qui a des ressources abondantes et variées ; de ce qui est pourvu de grandes qualités.	2. Qui annonce ou suppose la richesse, des moyens financiers importants.	2. Qui annonce ou suppose la richesse, des moyens financiers importants.
3. Qui est d'un grand prix.	3. Qui contient de nombreux éléments ou des éléments importants en abondance.	3. Qui contient de nombreux éléments ou des éléments importants en abondance.

Tableau 1. L'entrée RICHE dans trois dictionnaires

D'une manière analogue, dans le *Robert Culturel*, nous trouvons la citation suivante :

« Ah, que ne suis-je riche pour venir en aide au pauvre que je suis. » (Tristan Bernard)

L'interprétation de cet énoncé amène à penser que son auteur accepte une sorte d'axiome du genre : *si j'étais riche, je pourrais aider*. Aucune des définitions des dictionnaires cités ne laisse prédire que le mot *riche* établit une relation entre la possession et la capacité à aider, capacité qui n'est pas non plus (hélas !) garantie par la possession. L'exemple utilisé permet au lecteur de compenser ce manque, en suggérant qu'une telle relation peut parfois être établie.

Enfin, les dictionnaires étudiés présentent de multiples cooccurrences de l'adjectif en question avec des substantifs pour lesquelles ils proposent des synonymes, des explications ou des renvois :

De riches étoffes (syn. *magnifiques*) (*Larousse*)

Homme extrêmement riche, puissamment riche (= riche à millions, riche comme Crésus, cousu d'or.) (*Larousse*)

Aliment riche \Rightarrow *nourrissant*. (*Robert Culturel*)

Les gens riches, assez riches \Rightarrow *aisé*. (*Petit Robert*)

Comme précédemment, le fait que ces cooccurrences soient mentionnées indique que leur interprétation est considérée par les auteurs comme révélatrice de propriétés sémantiques du mot. Mais, là encore, ni la synonymie de ces synonymes ni l'explicativité de ces explications ne découlent des définitions proposées pour cette entrée.

À partir de l'étude de cette entrée, on peut formuler, en généralisant, un premier défaut de systématisation.

Pour certaines caractéristiques sémantiques des mots qu'ils décrivent, les dictionnaires étudiés les présentent de manière systématique, dans une définition ; pour d'autres caractéristiques sémantiques, les auteurs les évoquent, de manière implicite, en se fondant sur l'intuition que les lecteurs peuvent avoir à la lecture d'un exemple d'utilisation. Les deux classes de caractéristiques sont pourtant vraisemblablement considérées comme constitutives de la signification du mot, puisqu'elles apparaissent dans l'article.

1.4 GESTION DE LA SUBJECTIVITÉ

Le *Robert Culturel* permet d'illustrer un deuxième défaut de systématique. On y trouve l'énoncé suivant (toujours à propos de l'entrée « riche ») :

1.a Adj. Qui a de la fortune, possède des biens, de l'argent en abondance ou plus que la majorité des membres de sa communauté. (*Robert Culturel*)

Ainsi, par exemple, selon la définition que propose le *Robert Culturel*, un individu, Jean, ayant un salaire de 1 500 euros par mois et vivant dans une communauté dans laquelle le salaire de la majorité ne dépasse pas les 1 000 euros, sera considéré comme *riche* par un locuteur de cette communauté, puisque il « possède plus que la majorité des membres de sa communauté ». En revanche, du point de vue de quelqu'un venant d'une autre communauté, dont le salaire moyen est de 3 000 euros, Jean ne sera plus considéré comme riche. « Plus que la majorité » est un élément subjectif qui dépend du point de vue du locuteur. Cette partie de la définition peut être considérée comme une indication de subjectivité, mais qui reste implicite dans la définition. On voit ici que les auteurs du dictionnaire, en utilisant dans la description « plus que la majorité », ont repéré un phénomène subjectif lié à la description du mot, dont ils pensent devoir rendre compte. Ce que vise le *Robert* est probablement quelque chose comme « plus que ce que le locuteur considère comme la majorité de ce que le locuteur considère comme les membres de ce que le locuteur considère comme sa communauté » : une telle description serait, bien entendu, difficilement acceptable ; il n'en reste pas moins que l'expression utilisée à la place est inexacte : sous une apparence objective, la définition « est riche celui qui possède plus que la majorité de sa communauté » est, comme on vient de le voir, un énoncé subjectif à cause de la procédure de détermination des entités correspondant à « la majorité », « les membres », « la communauté ».

Le fait de devoir rendre compte d'un phénomène subjectif n'oblige pourtant pas à adopter un métadiscours subjectif ; on l'a vu dans notre tentative (inten-

tionnellement) maladroite de reformuler ce que nous pensons que l'auteur de l'article *riche* voulait dire. La définition proposée par le *Robert* donne ainsi l'impression que la subjectivité de l'évaluation qu'impose le mot *riche* était passée du locuteur à l'auteur de l'article. Bien entendu, ce glissement de subjectivité n'est pas systématique, et les auteurs des articles parviennent habituellement à faire comprendre à leurs lecteurs qu'un mot induit un jugement subjectif du locuteur.

La manière dont la subjectivité induite par les mots de la langue est décrite, subjectivité qui caractérise toute communication humaine, constitue donc un autre défaut de systématisation des dictionnaires étudiés.

1.5 ÉNONCÉS BIZARRES

Après avoir attentivement étudié les définitions et les exemples proposés dans n'importe lequel de ces trois dictionnaires, rien n'empêcherait un locuteur (pas très francophone) de déclarer, à propos d'un nouveau-né : « ce bébé est riche ». En effet, d'après les descriptions des dictionnaires étudiés, un tel énoncé ne devrait pas produire d'effet particulier, puisqu'un bébé peut très bien posséder plus que la majorité de sa communauté. Or pour les locuteurs francophones, cet énoncé provoque un effet bizarre (v. Bruxelles *et al.* 1995). Certes, le caractère bizarre d'un énoncé n'interdit pas son utilisation, mais une description sémantique convenable, dans un dictionnaire de langue, devrait permettre de prévoir de tels effets.

Ainsi, un troisième défaut de systématisation des dictionnaires est que la description des mots ne rend pas compte de phénomènes de restriction sémantique, comme le fait que certains énoncés grammaticalement corrects, peuvent paraître sémantiquement bizarres, bien que ces restrictions soient caractéristiques des mots.

Nous venons de voir que ce qui est systématique, dans la description des dictionnaires, c'est l'aspect dénotationnel (objectif), tandis que les aspects subjectifs des mots de la langue, eux, ne sont pas décrits de manière systématique. Les auteurs des dictionnaires (lexicographes, mais aussi, en amont, lexicologues et sémanticiens) décrivent les aspects subjectifs *apparaissant dans les énoncés*, mais ils ne font pas l'abstraction et l'abduction nécessaires pour décrire ces aspects en langue.

Nous proposons un moyen de combler ce déficit (traiter de manière systématique la subjectivité que les mots d'une langue induisent), en décrivant ces aspects de la subjectivité que nous appelons les *présupposés idéologiques* ou les *points de vue implicites cristallisés* dans les mots. Nous allons montrer qu'ils peuvent être décrits comme des *points de vue* que les mots de la langue imposent comme nécessaires à l'interprétation, et que leur prise en compte permet de résoudre les difficultés des trois types examinés ci-dessus. On verra en particulier, au paragraphe 2.3, comment décrire de manière systématique ces points de vue cristallisés dans les mots.

2. LES INDICATIONS SUR LES PRÉSUPPOSÉS IDÉOLOGIQUES CRISTALLISÉS

Une des difficultés à surmonter qui découle de la réflexion précédente est que la langue, en tant qu'objet à décrire par le sémanticien, n'est pas une entité directement observable – en entendant par *observation* la perception d'une entité par

un ou plusieurs organes de l'appareil sensoriel humain. Les seuls objets qui peuvent être observés sont les discours. C'est l'observation des discours qui permet au sémanticien d'observer la langue de manière indirecte. Par ailleurs, si l'on admet, avec une bonne partie des sémanticiens, que

la sémantique a pour objectif de décrire les contraintes que les phrases des langues humaines imposent aux locuteurs et interlocuteurs de ces langues dans le processus qui les conduit à construire un sens pour leurs énoncés (Raccah 2002 : 243),

alors, on est amené à admettre aussi que les propriétés des discours, et, en particulier, les sens des mots dans des discours ne constituent pas l'objet de description de la sémantique, mais seulement le point de départ empirique, à partir duquel la description sémantique est élaborée.

C'est dans ce sens que le dictionnaire se situe, en principe, entre les deux entités (langue et discours) : il est censé être du métadiscours du linguiste, c'est-à-dire des descriptions de la langue, tout en permettant aux utilisateurs des dictionnaires de produire ou de comprendre des discours. Or, la base empirique des descriptions que le dictionnaire fournit ne peut relever que des discours : les descriptions *en langue* sont nécessairement le résultat d'un raisonnement abductif qui conduit à formuler des hypothèses (concernant la langue) destinées à expliquer l'effet des mots sur l'interprétation des discours. Il en résulte aussi que les exemples et illustrations ne peuvent pas appuyer directement la description de la signification des mots, mais seulement les interprétations que cette signification autorise. Prenant cette remarque en considération, la question que nous nous posons est celle de savoir comment ces métadiscours doivent être construits et organisés ; que doivent contenir ces descriptions pour permettre un usage adéquat des mots ? Pour y répondre, compte tenu des insuffisances que nous avons trouvées dans les dictionnaires et des conclusions que nous en avons tirées, nous examinons d'abord comment des propriétés sémantiques, objectives et indépendantes des situations, attribuées aux unités de langue, peuvent rendre compte des usages subjectifs de ces unités en situations. À partir des conclusions de cet examen, nous proposons des éléments d'une grille de description sémantique, inspirée du cadre théorique de l'*argumentation dans la langue* (v. Ducrot 1983, Raccah 1987, Ducrot 1988 et Bruxelles *et al.* 1995), et dont nous montrons qu'elle constitue l'architecture d'un modèle à la fois descriptif de la signification et explicatif du sens. Nous verrons que son application à la description lexicale permet d'éviter les problèmes que nous avons mentionnés dans notre première partie.

2.1 DESCRIPTION SÉMANTIQUE, CONTRAINTES ET ORIENTATIONS ARGUMENTATIVES

Nous avons fait allusion au fait que les unités de langue ne sont pas porteuses de *sens*, mais de *contraintes* sur la construction du sens que leurs énoncés provoquent en situation. Or, s'il est indéniablement faux que tout énoncé constitue une argumentation, il n'en reste pas moins vrai (tout aussi indéniablement...) que toute phrase, quelle qu'elle soit, peut être énoncée pour constituer un argument à l'appui d'une conclusion. Une conséquence importante de ce fait est que les phrases imposent des contraintes sur les orientations argumentatives éventuelles de leurs énoncés. Il résulte de cela que la description sémantique des unités de langue doit comporter des contraintes sur ces orientations argumentatives.

Mais la distinction fondamentale entre les unités de langue (comme les *phrases*) et les unités de discours (comme les *énoncés*) impose une distinction corrélative entre :

- *mot-d'énoncé*, qui a un *sens* pour un interprète dans une situation (*sens* caractérisé par son rôle dans le *sens* de l'énoncé global),

et

- *mot-de-phrase*, qui a une *signification* (laquelle consiste en l'ensemble des *contraintes* que ce *mot-de-phrase* impose à la *construction du sens de l'énoncé*, dans les diverses situations).

Il résulte de la nécessité d'une telle distinction que la description, même rigoureusement motivée, du *sens d'un mot* (ou d'un groupe de mots) dans un énoncé ne constitue pas la description de la signification du mot-de-phrase qui lui est sous-jacent, même si la description de la signification d'un mot-de-phrase doit, bien entendu, rendre compte des différents sens des mots-d'énoncé auxquels il est sous-jacent⁹.

Or, un dictionnaire décrit des unités de langue, c'est-à-dire, pour ce qui nous concerne ici, des mots-de-phrase : il doit donc faire apparaître, pour chaque mot, toutes les contraintes que ce mot impose à la construction du sens des énoncés de toute phrase qui le contient, et ce, dans toute situation. Les dictionnaires mentionnés, en ne généralisant que les contraintes relatives à la référence, sont obligés, pour ne pas abandonner toute adéquation empirique, de traiter les autres types de contraintes au cas par cas et en fonction des situations, au moyen d'exemples dont on espère que le lecteur saura tirer (intuitivement) une règle générale que le lexicographe n'a pas lui-même su expliciter.

Ainsi, bien que les mots-de-phrase contraignent l'orientation argumentative des énoncés des phrases les contenant, les dictionnaires n'en systématisent pas la description : ces contraintes, lorsqu'un article y fait allusion, n'apparaissent qu'indirectement, à l'occasion d'exemples¹⁰.

2.2 ORIENTATION ARGUMENTATIVE, POINT DE VUE ET IDÉOLOGIE¹¹

Plusieurs des concepts et des notions utilisés ici s'appuient sur des notions initialement introduites par Oswald Ducrot, retravaillées par Ducrot lui-même, par Jean-Claude Anscombre, et par bien d'autres. Les signifiants utilisés par les uns ou les autres ne recouvrent plus toujours les mêmes notions, et, encore moins, les mêmes concepts. C'est en nous appuyant sur l'analyse sémantique d'une structure de phrase, [A *mais* B], que nous précisons nos concepts et exposons les liens entre les notions d'*orientation argumentative*, de *point de vue* et d'*idéo-*

9 Ainsi, l'analyse des discours ne suffit pas à la sémantique, même si la sémantique ne peut pas se passer d'analyse de discours.

10 C'est pourquoi ces exemples, qui ne devraient être que des illustrations des définitions, en sont des compléments nécessaires mais partiels.

11 Nous utilisons l'expression « orientation argumentative » dans le sens que l'École ducrotienne lui donne habituellement ; le syntagme « point de vue » est utilisé dans un sens que l'on précisera plus bas et qui peut se décrire comme une abstraction commune du concept d'*orientation argumentative* et de celui d'*énonciateur polyphonique* ; le terme « idéologie » est utilisé dans un sens technique, défini dans Raccah (2002), et désigne tout point de vue qu'il est nécessaire d'admettre (même provisoirement) pour comprendre un énoncé ou un discours.

logie. Nous montrerons en outre pourquoi ces notions liées à l'interprétation des énoncés sont également importantes pour la description de la signification des mots-de-phrases. Le choix d'un connecteur dont la description argumentative est assez connue permet d'introduire la terminologie utilisée ici, et de préciser les concepts mis en œuvre, en soulignant leur originalité et leur efficacité¹².

Pour comprendre un énoncé d'une phrase de structure [A *mais* B] renvoyant à une situation S, il est nécessaire de trouver, dans S, des raisons pour admettre qu'un énoncé de B renvoyant à cette situation viserait des conclusions argumentatives opposées à celles que viserait un énoncé de A dans cette même situation. Ainsi, dans une situation dans laquelle un intervenant propose à un autre de l'accompagner pour une promenade, une réponse comme :

(1) *Il fait chaud mais j'ai un peu mal au dos*

peut être comprise comme un refus de la proposition, refus que l'on peut analyser comme suit :

- a. le temps qu'il fait, qualifié de « chaud », est présenté comme un argument favorable à la promenade ;
- b. le fait que le locuteur se présente comme ayant un peu mal au dos constitue un argument défavorable à la promenade ;
- c. le locuteur choisit de décider en suivant le second argument plutôt que le premier.¹³

Les contenus des trois éléments d'analyse précédents dépendent de la situation à laquelle l'interprétation fait renvoyer l'énoncé, telle que l'interprète la conçoit. En effet, un énoncé de cette même phrase peut aussi être interprété comme une acceptation de la proposition de promenade, acceptation que l'on peut analyser comme suit :

- a'. le temps qu'il fait, qualifié de « chaud », est présenté comme un argument défavorable à la promenade ;
- b'. le fait que le locuteur se présente comme ayant un peu mal au dos constitue un argument favorable à la promenade ;
- c'. le locuteur choisit de décider en suivant le second argument plutôt que le premier.

On constate que si les points de vue sont opposés, la structure de l'analyse, elle, ne dépend que de la présence de « mais » dans la phrase, indépendamment des situations d'interprétation¹⁴. De manière plus précise, (a) s'appuie sur l'idée que

Ta // Plus il fait chaud, plus on est enclin à se promener //

et (b) s'appuie sur l'idée que

12 Voir Raccah (1987) pour une analyse détaillée de *but* (angl.), ou Raccah (2002) pour une analyse de *mais*. Ces analyses reprennent et complètent celle qui a été, à l'origine, proposée par Ducrot (1972 : 128-129), utilisée depuis par bien des auteurs.

13 On remarquera que, si les deux membres sont intervertis, l'énoncé devenant :

(1') J'ai un peu mal au dos mais il fait chaud

la réponse constitue, dans cette même situation, une acceptation de la proposition de promenade : le locuteur choisit, encore une fois, de décider en suivant le second argument plutôt que le premier.

14 Voir aussi Raccah (1987) pour la mise en évidence d'un phénomène analogue à propos du *but* anglais.

Tb // Plus on a mal au dos, moins on est enclin à se promener //

tandis que (a') s'appuie sur l'idée que

Ta' // Plus il fait chaud, moins on est enclin à se promener //

et (b') s'appuie sur l'idée que

Tb' // Plus on a mal au dos, plus on est enclin à se promener //

En généralisant cette discussion, on peut formuler l'hypothèse descriptive suivante, concernant toute phrase [A *mais* B] de la langue française contenant le mot-de-langue *mais* :

D₁[A *mais* B] : Dans toute situation S dans laquelle un énoncé de la phrase [A *mais* B] est interprétable, un énoncé de la phrase B devrait être compris comme orienté vers une conclusion opposée à la conclusion vers laquelle serait orienté un énoncé de la phrase A dans cette même situation.

D₂[A *mais* B] : Dans toute situation S dans laquelle un énoncé de la phrase [A *mais* B] est interprétable, l'orientation argumentative de cet énoncé de [A *mais* B] est celle qu'aurait un énoncé de B dans cette même situation.¹⁵

Ainsi, l'élément D₂ de la description de [A *mais* B] peut être conçu comme une contrainte sur les points de vue que les énoncés de [A *mais* B] peuvent servir à construire, contrainte selon laquelle ces points de vue sont les mêmes que ceux que les énoncés de B peuvent servir à construire. La contrainte correspondant à D₁ est, quant à elle, une contrainte *préalable* : pour pouvoir interpréter un énoncé de [A *mais* B] dans une situation S, il est nécessaire d'*avoir admis* que le point de vue que l'on peut construire en interprétant un énoncé de B dans S est l'opposé de celui que l'on peut construire en interprétant un énoncé de A dans S. Pour s'en convaincre, il suffit d'instancier la structure [A *mais* B] avec deux phrases A et B qui soient telles que, dans les situations habituelles, les énoncés de A et de B n'appellent pas des points de vue opposés : on se rend compte alors qu'un énoncé de [A *mais* B] n'est interprétable dans aucune de ces situations. Si l'on tient malgré tout à interpréter l'énoncé, pour y parvenir, un effort d'imagination est nécessaire : il faut chercher des situations dans lesquelles un énoncé de A et un énoncé de B sont susceptibles d'appeler des points de vue opposés. C'est ce parcours interprétatif qui est mis à l'œuvre pour pouvoir comprendre :

(2) *Il pleut mais j'ai du travail*

en réponse à une proposition de promenade.

Si l'on accepte une caractérisation non idéologique de l'idéologie, selon laquelle relève de l'idéologie tout ce qu'un discours oblige (ou se présente comme obligeant) à admettre¹⁶, des contraintes comme D₁, préalables à l'interprétation, constituent le cadre idéologique du discours qui les met en œuvre. Ainsi, pour interpréter un énoncé comme :

15 Le lecteur averti reconnaîtra une parenté entre cette description et la description originale proposée par Ducrot (1972 : 128-129), à ceci près que, d'une part, notre description n'identifie plus les énoncés à des propositions et, d'autre part, elle relativise explicitement et systématiquement la valeur d'opposition à la situation d'interprétation.

16 Une caractérisation idéologique de l'idéologie spécifierait que, pour relever de l'idéologie, ces conceptions doivent en outre être fausses ou injustifiées...

(3) *Paul est ministre mais il est honnête*

il faut admettre (au moins pendant le temps nécessaire à la compréhension) que les deux points de vue exprimés par l'un et l'autre des membres de l'énoncé s'opposent, c'est-à-dire que les ministres sont, en règle générale, malhonnêtes ou peu honnêtes¹⁷.

3. QU'EST-CE QU'UN POINT DE VUE LEXICAL ?

Au paragraphe précédent, nous avons vu que la présence de *mais* dans une phrase introduit des contraintes sur les rapports entre les points de vue des énoncés de cette phrase et des éventuels énoncés de parties de cette phrase. Mais les contraintes sémantiques sur les points de vue ne se limitent pas à des rapports entre points de vue. Ainsi, dans l'exemple (3), la présence du mot *honnête* suggère un point de vue positif sur les qualités morales de la personne à propos de laquelle l'énoncé discourt. Il est intéressant de remarquer que, quelle que soit l'idéologie du locuteur, l'utilisation du mot *honnête* impose un point de vue positif : même un mafieux exprime un point de vue positif sur Paul en disant :

(4) *Paul est honnête.*

De tels mots, qui imposent un point de vue positif *dans toute situation*, sont appelés « euphoriques ». D'autres mots imposent un point de vue négatif dans toute situation : ils sont dits « dysphoriques » (c'est le cas, par exemple, de *malhonnête* ou de *minable*).

D'autres mots, enfin, imposent une évaluation plus subtile : ils obligent à voir une entité ou une qualité à travers le point de vue à travers lequel on verrait une autre entité ou qualité. C'est le cas, par exemple, du mot *riche*, que nous avons décrit au § 1.2. Nous avons vu, en effet, que les différents aspects des sens des énoncés faisant intervenir le mot *riche*, et utilisés par les dictionnaires pour en illustrer la signification, ne pouvaient pas être prédits par une description dénotationnelle, laquelle ne permettait pas d'inscrire, dans la signification du mot-de-langue, l'idée que, pour utiliser ce mot, il est nécessaire de voir la possession comme source de pouvoir. Or, c'est cette idée qui est sous-jacente à toutes les illustrations que les dictionnaires utilisent pour pallier l'incomplétude de la définition dénotationnelle. Pour fournir une description sémantique adéquate du mot-de-langue *riche*, il faut donc indiquer que toute phrase l'utilisant oblige à voir la possession comme source d'un certain pouvoir. Ce sont les indications de ce type qui constituent les *points de vue lexicaux*, les *présupposés idéologiques cristallisés* dans le lexique.

Le mot *riche* est un cas prototypique qui illustre bien la nécessité de décrire ces points de vue cristallisés dans les mots-de-langue, car les présupposés idéologiques qu'il cristallise sont faciles à mettre en lumière (§ 1.2 ci-dessus). Son analyse a permis de révéler un phénomène dont tout laisse à penser qu'il est beaucoup plus général qu'une simple curiosité : la cristallisation de points de vue dans le lexique des langues.

17 Le lecteur averti reconnaîtra l'emprunt au célèbre exemple de George Lakoff : *John is a republican but he is honest* (Lakoff 1971 : 67). Il reconnaîtra aussi que ce qui en est dit ici n'a pas grand-chose à voir avec ce pourquoi Lakoff utilisait cet exemple.

3.1 POUR UNE DESCRIPTION SÉMANTIQUE DES POINTS DE VUE LEXICAUX

Nous avons vu que les descriptions lexicographiques actuelles présentaient des faiblesses du point de vue sémantique. Nous avons vu comment ces faiblesses se manifestaient et la nature des aspects sémantiques qui manquent à ces descriptions. Nous avons ensuite rattaché ces analyses à une réflexion théorique conduisant à un cahier des charges de la description sémantique. Il nous faut maintenant montrer la manière dont ce cahier des charges peut être satisfait : expliciter un modèle descriptif de la sémantique lexicale permettant de décrire ces *présupposés idéologiques cristallisés*, à propos desquels nous venons de voir, d'une part, que les dictionnaires en suggèrent des descriptions non systématiques et partielles et, d'autre part, qu'une sémantique adéquate devrait les décrire de manière systématique.

Pour ce faire, nous mettons à profit les relations que nous venons d'exposer entre orientation argumentative, point de vue et idéologie. Le raisonnement conduisant au modèle descriptif peut être schématisé de la manière suivante :

1. les *présupposés idéologiques cristallisés* sont des points de vue que les mots-de-langue imposent comme conditions préalables à toute interprétation des énoncés qui les utilisent ;
2. les points de vue imposés par les discours ou par des segments de discours sont exprimables en termes d'orientations argumentatives de ces discours ou de ces segments de discours ;
3. les orientations argumentatives de segments de discours s'expriment en termes de champs topiques (v. Raccah 1990 et Bruxelles *et al.* 1995) ;
4. les *présupposés idéologiques cristallisés* sont donc exprimables en termes de champs topiques lexicalisés.

Raccah (1990) propose une définition récursive du concept de « champ topique », définition qui permet de caractériser l'idée que le point de vue que l'on a sur une entité peut constituer un principe de valuation qui déterminera, à son tour, le point de vue que l'on aura sur une autre entité (v. Bruxelles *et al.* 1995). Tout champ topique peut donc être vu à la fois comme exprimant un point de vue et comme exprimant un principe d'argumentation : ce phénomène permet d'associer, de manière canonique (garantissant l'existence et l'unicité), un topos à tout champ topique. Un champ topique lexical peut alors jouer le rôle de contrainte sur la construction de points de vue.

Dans cette définition, les champs topiques sont des chaînes de couples composés de champs conceptuels et de principes d'évaluation, lesquels peuvent être eux-mêmes des champs topiques (d'où la récursivité). Pour le champ topique associé au mot *riche*, on a :

$$CT_{riche} = \langle \text{POSSESSION, } pouvoir \rangle$$

où *pouvoir* est lui-même un champ topique, que l'on peut décrire comme :

$$CT_{pouvoir} = \langle \text{CAPACITÉ_D'ACTION, } bien \rangle$$

ou bien :

$$CT_{pouvoir} = \langle \text{CAPACITÉ_D'ACTION, } mal \rangle$$

On remarquera que ce modèle descriptif prédit (correctement) deux points de vue opposés lexicalisés dans le mot-de-langue *riche*.

3.2 EXEMPLE : *JOUER*

Après avoir brièvement présenté les outils qui permettent aux dictionnaires de systématiser la description explicite des points de vue lexicalisés dans les mots, il convient de montrer la manière dont ces outils s'utilisent effectivement.

Nous prenons ici l'exemple des mots *jeu*, *jouer*¹⁸.

La question que nous nous posons est donc celle-ci : Quel est ou quels sont les points de vue qu'un locuteur est obligé d'adopter pour pouvoir employer un mot comme *jeu* ? Quel est le point de vue à partir duquel on est obligé de considérer une activité (le référent) pour pouvoir l'appeler *jeu* ?

Ou encore, du point de vue des dictionnaires : Parmi les présupposés idéologiques que les dictionnaires indiquent sous forme d'exemples, de synonymies, etc., lesquels doivent être considérés comme cristallisés dans le mot *jeu*, c'est-à-dire comme faisant partie de sa signification et lesquels doivent être considérés comme étant des présupposés idéologiques non cristallisés, des associations culturelles ?

Une première idée qui vient à l'esprit quand on pense au mot *jeu* est celle de l'amusement, du divertissement. C'est par le point de vue du divertissement que le *Micro-Robert* (1988) débute l'article *jeu* :

Activité physique ou mentale purement gratuite, qui n'a, dans la conscience de la personne qui s'y livre, d'autre but que le plaisir qu'elle procure.

Dans l'article *jouer* (*ibid.*), nous trouvons l'indication d'une synonymie entre *jouer* et *s'amuser* :

I. V. intr. 1. Se livrer au jeu ⇒ s'amuser. II. Suivi d'une prép. 2. jouer avec qqch. Petite fille qui joue avec sa poupée. ⇒ s'amuser.

D'où un premier candidat de présupposé idéologique (on verra plus tard s'il est cristallisé ou non) des mots *jeu*, *jouer*, exprimé sous forme de point de vue : on emploie ce mot quand on considère l'activité en question du point de vue de l'amusement, du divertissement qu'elle procure :

CT1 : <ACTIVITÉ, *amusement*>

Une deuxième idée que l'on associe souvent au mot *jeu* est celle de la « facilité ». C'est cette idée qui s'exprime dans l'expression « c'est un jeu d'enfant », considérée comme synonyme de « c'est facile ». C'est ce que l'on voit apparaître dans la description donnée par le *Micro-Robert* (*ibid.*) :

V. Loc. (dans lesquelles *jeu* s'applique à des actions, des activités, des affaires). *C'est un jeu d'enfant*, la chose est très facile.

L'article *jeu* (et pas seulement la locution mentionnée), toujours dans le *Micro-Robert* évoque la même idée, en y associant celle de gratuité, futilité, etc. :

Activité qui présente un ou plusieurs caractères du jeu (gratuité, futilité, bénignité, facilité).

C'est cette même interprétation qui paraît dans l'expression « c'est un jeu pour X de faire Y ». Mais, comme on le remarquera, l'idée de facilité n'est pas exprimée de manière explicite, c'est dans l'interprétation qu'elle apparaît.

¹⁸ Dans cet exemple, la catégorie syntaxique n'est pas pertinente du point de vue sémantique. Voir Nemo (2002) à propos du noyau sémantique commun que constituent les morphèmes.

Néanmoins, en attendant de tester son acceptabilité, formulons l'hypothèse que dans *jeu*, il y a le point de vue de facilité, ce qui donne :

CT2 : <ACTIVITÉ, *facilité*>

Une troisième idée associée au mot *jeu*, qui apparaît à travers l'examen des discours, est celle de l'absence du sérieux dans l'activité décrite. *Jeu* s'oppose à *activité sérieuse*. Les locuteurs, les narrateurs décrivent des activités, des comportements par l'expression *jeu* quand ils les voient comme une activité non sérieuse.

— Et pourquoi ?... *c'est un jeu pour vous, une chose sérieuse pour moi...* je vous ai provoqué... j'ai excité chez vous... un petit sentiment... j'ai été imprudente... je viens de vous dire mille paroles sans suite... les premières qui me sont venues à vous dire... allez ! Ce n'est pas grand-chose chez vous tout cela... et ça ne laissera pas grande trace... il vaut mieux qu'il n'y ait rien entre nous... (Goncourt 1860 : 76)

D'où notre troisième candidat :

CT3 : <ACTIVITÉ, *légèreté*>

On trouve un emploi intéressant du mot *jeu* dans certains contextes de dialogue. Il s'agit de l'emploi où un locuteur utilise l'énoncé « C'est un jeu » pour calmer son interlocuteur, pour l'amener à ne pas avoir peur, comme dans la citation suivante :

[...] un étroit et obscur couloir sans issue, nous allons piétiner sans fin, enfermés avec elle dans ce labyrinthe sombre et clos, tournant en rond... Mais *n'ayez pas peur...* *C'est un jeu*, vous le savez bien... Aucun de nous ne risque rien. Le cœur se serre délicieusement, on a envie de crier, comme sur les montagnes russes quand le wagonnet descend, on rit [...] (Sarraute 1959 : 30-31)

Cet emploi du mot *jeu* nous amène au quatrième candidat :

CT4 : <ACTIVITÉ, *tranquillité*>

Au lieu de nous limiter à une liste de points de vue, que l'on pourrait allonger en variant les situations, nous pouvons tenter d'approfondir la réflexion et de nous demander si la possibilité d'employer le mot *jeu* pour amener l'interlocuteur à ne pas prendre en compte la peur ou le sérieux, ou à voir l'activité comme amusante ou facile, si cette possibilité n'est pas une conséquence de la cristallisation d'une croyance plus élémentaire : celle qui associe l'absence de conséquences aux activités pouvant être considérées comme *jeu*. Cette pensée nous paraît d'autant plus fondée que le *Micro-Robert* mentionne cet effet de sens :

Ce n'est qu'un jeu, cela ne tire pas à conséquence.

D'où notre nouvelle proposition comme présupposé idéologique de *jeu* :

CT5 : <ACTIVITÉ, absence_de_conséquences>

À partir des descriptions existantes dans les dictionnaires, à partir de l'observation des discours, on arrive donc à cinq candidats pour le champ topique lexical. Le mot *jeu* pourrait avoir, d'après ces hypothèses, comme point de vue lexical :

- l'amusement :

CT1 : <ACTIVITÉ, *amusement*>

- la facilité :

CT2 : <ACTIVITÉ, *facilité*>

- l'absence de sérieux :

CT3 : <ACTIVITÉ, *légèreté*>

- l'absence de peur :

CT4 : <ACTIVITÉ, *tranquillité*>

- l'absence de conséquences :

CT5 : <ACTIVITÉ, *absence_de_conséquences*>

Pour décider lequel ou lesquels de ces présupposés idéologiques sont cristallisés dans le mot, nous procédons selon une démarche conforme au paradigme poppérien de la scientificité : nous avons tenté de les réfuter au moyen de tests linguistiques. Ces tests linguistiques, dont l'élaboration a été l'objectif de la thèse d'un des auteurs ¹⁹, sont basés sur la notion d'énoncé sémantiquement paradoxal et non interprétable : un énoncé qui *présuppose* le point de vue contraire à celui qui est lexicalisé dans un des mots de l'énoncé. Pour obtenir un tel énoncé de test, nous utilisons l'idée, développée ci-dessus, selon laquelle les connecteurs et les opérateurs contraignent les points de vue qui apparaissent dans les énoncés. Certains connecteurs *présupposent* une relation argumentative entre les segments qu'ils relient. Ainsi, si on construit un énoncé dans lequel figure :

- un mot M (en présupposant son point de vue lexical) ;
- la formulation du contraire du point de vue supposé lexical L du mot M ;
- et un connecteur qui présuppose une relation argumentative entre les segments reliés (M et non-L) ;

alors, on devrait obtenir un énoncé non interprétable parce que le locuteur d'un tel énoncé devraient prendre en charges deux présupposés contraires.

À la suite de ce type de tests, il s'avère que, parmi les cinq candidats, quatre ne sont pas cristallisés dans le mot *jeu* : les points de vue de l'*amusement*, de la *facilité*, du *non sérieux*, et de l'*absence de peur*. Ces présupposés idéologiques sont culturels mais pas linguistiques. En effet, on peut avoir des énoncés dans lesquels ces présupposés idéologiques sont annulés sans pour autant donner un énoncé in-interprétable :

C'est un jeu, il n'est pourtant pas sérieux.

C'est un jeu, n'aie pourtant pas peur.

semblent être des énoncés in-interprétables à première vue, mais en fait, on trouve facilement des situations dans lesquelles on arrive à les interpréter : en admettant l'hypothèse que le locuteur du premier énoncé n'aime que les jeux sérieux, ou que le destinataire du deuxième est une personne qui est intimidée par l'idée de jouer.

Seul le point de vue de l'absence des conséquences résiste aux tests : il est impossible d'utiliser le mot « jeu » sans adopter le point de vue en question sur

¹⁹ E. Chmelik, *Idéologie dans les mots. Contribution à une description topique du lexique justifiée par des tests sémantiques. Application à la langue hongroise*, Université de Limoges, 2007.

l'activité pour laquelle nous utilisons le mot ; si on essaye de le faire, on a des énoncés in-interprétables :

**C'est un jeu, il n'a pourtant pas de conséquences.*

nous semble correspondre, en effet, à un énoncé in-interprétable, même en construisant des hypothèses supplémentaires sur son locuteur.

Pour conclure cette étude sémantique sur les mots *jeu* et *jouer*, ce qui nous semble nécessaire, c'est de marquer, dans les dictionnaires, de manière systématique les présupposés idéologiques cristallisés dans ces mots : activité vue du point de vue des conséquences, car ces présupposés idéologiques ne peuvent pas être annulés dans les discours... On les distingue ainsi de ceux qui ne sont que culturels (comme l'amusement, la facilité) qui peuvent ne pas être présents dans les discours.

Plus généralement, nous avons montré que les dictionnaires que nous avons étudiés proposent une description explicite et systématique des aspects dénotationnels de la signification des mots-de-langue, mais n'évoquent que de manière implicite et allusive ses aspects implicites et subjectifs, jugés pourtant suffisamment importants, par les auteurs des dictionnaires, pour qu'ils illustrent ces aspects dans les articles décrivant les mots. Nous avons suggéré que cet état de fait s'explique par une faiblesse des théories sémantiques sous-jacentes aux dictionnaires, faiblesse dont un des effets est que les relations entre points de vue subjectifs ne pouvaient pas être décrites objectivement et systématiquement. Nous proposons un modèle théorique, inspiré des travaux originaux de l'école ducrotienne de l'*argumentation dans la langue*, dans lequel la sémantique peut et doit décrire les contraintes que les unités de langue imposent aux points de vue que leurs énoncés expriment. Nous montrons enfin, à titre d'illustration, comment ce modèle théorique permet de décrire les points de vue cristallisés dans les mots *jeu*, *jouer* de la langue française.

Le modèle présenté, bien qu'inspiré des travaux originaux d'Oswald Ducrot, intègre d'autres préoccupations épistémologiques, méthodologiques et théoriques, qui ont été abondamment décrites ailleurs et qui font qu'il se distingue très nettement des nombreux travaux se réclamant de l'école ducrotienne. Ce n'est pas le rôle de cet article de reprendre, plus de quinze ans après les premières publications sur ces différences, les analyses des similitudes et des divergences entre objectifs et méthodes des différentes démarches s'inspirant, de près ou de loin, des travaux originaux de Ducrot.

Il convient de garder à l'esprit que notre (méta-)argumentation s'appuie sur l'analyse d'exemples et sur une illustration concernant un cas particulier : elle n'a donc pas valeur de *preuve formelle* de ce que nous avançons, ni de l'efficacité réelle de nos propositions. Ces preuves formelles, dont certains éléments ont été publiés et sont mentionnés dans cet article, doivent encore être rassemblées et complétées pour qu'elles puissent prendre une valeur paradigmatique. Un tel travail ne peut évidemment pas faire l'objet d'un simple article. En revanche, cet article peut suggérer des pistes à explorer pour approfondir les recherches et/ou pour trouver des dispositifs permettant de tester les différentes thèses qui ont été présentées. C'est l'effet que nous souhaitons obtenir.

BIBLIOGRAPHIE

- BRUXELLES Sylvie, DUCROT Oswald and RACCAH Pierre-Yves, 1995, "Argumentation and the lexical topical fields", *Journal of Pragmatics* 24 (1/2), p. 99-104.
- CHMELIK Erzsébet (sous presse), « Des outils sémantiques pour révéler les biais inscrits dans la langue », dans P.-Y. Raccah (éd.), *L'Inscription des biais socio-cognitifs dans la langue*, Paris, L'Harmattan.
- COLLINOT André et MAZIÈRE Francine, 1997, *Un prêt à parler : le dictionnaire*, Paris, Puf.
- DUBOIS Jean et DUBOIS Claude, 1971, *Introduction à la lexicographie : le dictionnaire*, Paris, Larousse.
- DUCROT Oswald, 1972, *Dire et ne pas dire. Principes de sémantique linguistique*, Paris, Hermann.
- DUCROT Oswald, 1980, *Les Mots du discours*, Paris, Minuit.
- DUCROT Oswald, 1983, « Opérateurs argumentatifs et visée argumentative », *Cahiers de linguistique française*, 5, p. 7-36.
- DUCROT Oswald, 1984, *Le Dire et le Dit*, Paris, Minuit.
- DUCROT Oswald, 1988, « Topoi et formes topiques », *Bulletin d'études de linguistique française*, 22 (Tokyo).
- GAUDIN François et GUESPIN Louis, 2000, *Initiation à la lexicologie française. De la néologie aux dictionnaires*, Bruxelles, Duculot.
- DE GONCOURT Edmond et Jules, 1860, *Charles Demailly*, Paris, Charpentier.
- KLEIBER Georges, 1994, *Nominales, essais de sémantique référentielle*, Paris, Armand Colin.
- KREMER Jean-Marc, 2008, « La petite histoire des grands dictionnaires ou À quoi sert la lexicographie ? » Dossier en ligne sur le site de la Fédération Nationale des Orthophonistes, dernière consultation : février 2008 : <http://www.orthophonistes.fr/upload/DO2-lexicographie.pdf>
- LAKOFF George, 1971, "Presuppositions and relative well-formedness", in D. D. Steinberg and L. A. Jakobovits (eds.), *Semantics: An Interdisciplinary Reader*, Cambridge, Cambridge University Press, p. 329-340.
- MARTIN-BERTHET Françoise et LEHMANN Alise, 2000, *Introduction à la lexicologie : Sémantique et morphologie*, Paris, Nathan.
- MATORÉ Georges, 1953, *La Méthode en lexicologie, domaine français*, Paris, Didier.
- MORTUREUX Marie-Françoise, 1997, *La Lexicologie entre langue et discours*, Paris, Sedes.
- NEMO François, 2002, "Morpheme semantics and the autonomy of morphology. The stable semantics of (apparently) unstable constructions", in M. Andronis, C. Ball, H. Elston and S. Neuvel (eds.), *CLS 37: The Panels. Papers from the 37th Meeting of the Chicago Linguistic Society. Vol. 2*. Chicago, Chicago Linguistic Society.
- QUEMADA Bernard, 1968, *Les Dictionnaires du français moderne (1539-1863)*, Paris, Didier.

- RACCAH Pierre-Yves, 1987, "Modelling argumentation and modelling with argumentation", *Argumentation*, 4, p. 447-483.
- RACCAH Pierre-Yves, 1990, « Signification, sens et connaissance : une approche topique », *Cahiers de Linguistique Française*, 11, p. 178-198.
- RACCAH Pierre-Yves, 1999. «¿Porqué los bebés españoles son más ricos que los bebés franceses?», *Quaderns de filologia. Estudis linguistics* 3, p. 1-17.
- RACCAH Pierre-Yves, 2002, « Lexique et idéologie. Les points de vue qui s'expriment avant qu'on ait parlé », dans M. Carel (éd.), *Les Facettes du dire. Hommage à Oswald Ducrot*, Paris, Kimé, p. 242-268.
- RACCAH Pierre-Yves, 2005, "What is an empirical theory of linguistic meaning a theory of?", in Z. Frajzyngier *et al.*, *Diversity and Language Theory Studies in Language*, Amsterdam and Philadelphia, Benjamins.
- REY Alain, 1977, *Le Lexique, images et modèles du dictionnaire à la lexicologie*, Paris, Armand Colin.
- REY-DEBOVE Josette, 1971, *Étude linguistique et sémiotique des dictionnaires français contemporains*, La Haye, Mouton.
- SARRAUTE Nathalie, [1959] 1993, *Le Planétarium*, Paris, Gallimard.
- SÉRASSET Gilles et POLGUÈRE Alain, 1997, « Outils pour lexicographes : application à la lexicologie explicative et combinatoire », *Actes de RIAO'97*, p. 701-708.

DICTIONNAIRES

- Larousse Lexis, Dictionnaire de la langue française*, 1999.
- Micro-Robert*, 1988.
- Petit Robert*, 2006.
- Robert Culturel*, 2005.

DEUXIÈME PARTIE

VALEURS D'EMPLOI

I.

MARQUE DIATOPIQUE

PROLÉGOMÈNES À UNE ÉTUDE DES MARQUES D'USAGE DANS LA LEXICOGRAPHIE DIFFÉRENTIELLE : LE CAS DU FRANÇAIS DE SUISSE ROMANDE

par Dorothée AQUINO-WEBER, Université de Neuchâtel,
et Christel NISSILLE, Université de Bâle

Résumé — Le cadre théorique permettant l'établissement des systèmes des marques d'usage est encore déficitaire dans la tradition lexicographique de l'espace francophone, et en particulier dans la lexicographie différentielle. Nous aimerions tenter de développer les références théoriques qui permettront de mieux caractériser la nature intrinsèque du français régional et l'outil théorique de sa description, et plus particulièrement du système de marquage du français régional. Il s'agira de comprendre de manière plus rigoureuse le « sentiment » du locuteur à l'égard de sa langue et les rapports qu'il entretient avec celle-ci par une évaluation de lexèmes identifiés comme régionaux. Cette évaluation sera basée sur l'élaboration d'une typologie de valorisation (ou de dévalorisation) des lexèmes concernés. Ceci devrait permettre la constitution d'une échelle de variation qui servira à dégager une « norme » régionale objectivée, et à mettre en lumière, dans l'axe de variation diatopique, l'existence de centres de référence et d'irradiation qui jusqu'à présent n'étaient que pressentis.

Mots clés — théorie lexicographique, marques d'usage, français régional, représentations linguistiques, norme.

Abstract — The theoretical frameworks underpinning dictionaries' labels and features are still dependent on mainland French-speaking lexicographical tradition. In this study the authors develop theoretical principles to better characterize the intrinsic nature of regional French. They also propose a theoretical tool for developing an improved system of labelling regional French. In particular, they suggest investigating and understanding the “feeling” of regional French speakers with regard to their own language, and advise that this feeling be considered when evaluating words identified as regional. The speaker's evaluation will enable lexicographers to develop a typology of valorised (or not valorised) words within the speaker's community and hence, the building of a sociolect variation scale. This tool could then be used to create an objective regional norm.

Keywords — theory of lexicography, labels and features, regional French, Swiss French, speakers linguistic representations, norm.

INTRODUCTION

Les réflexions que nous souhaitons présenter dans cet article s'inscrivent dans le cadre de la collaboration du Centre de dialectologie et d'étude du français régional de l'Université de Neuchâtel à la *Base de données lexicographiques pan-francophone*, plus commodément appelée BDLP¹, qui consiste en la mise en réseau des différentes descriptions lexicographiques différentielles des variétés régionales du français.

Durant le travail d'élaboration du volet suisse romand de la BDLP – la BDLPSuisse – de nombreuses questions se sont posées à nous. Elles nous ont amenées à repenser la manière dont il faudrait traiter les variantes, mais aussi à nous pencher sur le statut de l'emploi des lexèmes et leur classement géolinguistique. Toutes ces interrogations nous ont conduites à réfléchir sur le statut du régionalisme et de son marquage. Il nous a alors semblé nécessaire de poursuivre cette réflexion lexicographique et lexicologique de manière plus rigoureuse et de créer un « pont » entre la description lexicographique au sens strict et les recherches théoriques sur le français dans l'espace francophone. C'est la raison pour laquelle nous avons développé un projet de recherche en vue de le soumettre à divers organismes de financement.

L'objectif de cet article est double. Partant d'un constat de manque au niveau du cadre théorique pour l'établissement des systèmes de marques d'usage, nous souhaitons présenter les réflexions théoriques qui s'inscrivent dans ce projet de recherche ainsi que l'« outil » – le moyen d'appréhender la réception et la production de lexèmes régionaux – que nous avons élaboré pour tenter de caractériser le lexique régional de la Suisse romande².

Comme cet outil, qui s'apparente à une grille d'analyse, est actuellement en cours d'élaboration, nous n'avons pas encore récolté de données. Par conséquent, les différentes étapes que nous allons décrire et les résultats qui en découlent ne sont que théoriques. Par contre, pour illustrer notre réflexion, nous nous sommes basées sur un corpus partiel – essentiellement écrit – constitué de diverses citations extraites principalement de la littérature et de la presse, de textes administratifs et/ou officiels ou d'avis et études.

HYPOTHÈSES ET OBJECTIFS

Actuellement, dans la majeure partie des travaux de lexicographie différentielle, les lexiques régionaux sont abordés uniquement d'un point de vue géographique. En effet, ces ouvrages traitent la variation diatopique en la mettant toujours en rapport avec la norme traditionnelle.

1 La BDLP est un projet international actuellement dirigé par Claude Poirier (Université Laval, Québec) dont l'objectif est de décrire et de répertorier les variétés lexicales des différentes régions francophones à travers le monde (<http://www.tlfi.ulaval.ca/bdip/>, Poirier 2005).

2 Pour une description de la diversité interne de la Suisse romande et de son intégration dans l'Est (Savoie, Franche-Comté), voir Knecht 1985.

Pourtant, en cherchant à pousser la réflexion plus avant et à passer par delà la marque diatopique du lexème, qui n'est finalement qu'un simple critère de sélection de corpus, nous avons rencontré plusieurs difficultés. D'une part, il est important de définir précisément à quelle variété linguistique appartient le lexème. Par « variété linguistique », nous entendons « registre de langue », « type de discours », « canal de transmission », etc. D'autre part, lorsque le lexème s'inscrit dans un champ synonymique, il s'avère nécessaire de tenter de déterminer son sens ajouté – sa connotation qui, comme l'écrit Corbeil (1998 : 34), désigne une « information d'un autre ordre qui s'ajoute au sens et qui évoque des réalités d'une toute autre nature, ce qui différencie ce mot de ses parasyonymes » – par rapport aux autres occurrences qui font partie du même champ.

Nous postulons donc que le régionalisme est caractérisé par autre chose que sa seule marque diatopique et nous pensons qu'il s'agit désormais de dépasser le concept de marque régionale, objective et lexicographique, pour se pencher sur sa connotation, plus subjective.

A. Thibault et P. Knecht (1997), pour la constitution du *Dictionnaire suisse romand* (DSR), ont été contraints de se limiter à une démarche purement pragmatique enregistrant « ce qui se dit ou s'écrit, et, à l'exception des termes officiels, ne prescri[vant] pas ce qui doit se dire ou s'écrire » (Thibault 1997 : 13). En effet, en ce qui concerne plus particulièrement l'application de marqueurs à leur nomenclature, les auteurs de ce dictionnaire ont, comme l'explique Thibault (2000 : 77), eu recours à « une solution [...] de nature énonciative » :

le lexicographe doit tenter de séparer le plus clairement possible son discours à lui, que l'on peut qualifier d'« explicitant », du discours objet de son étude, lequel est « explicité ».

Ainsi, en privilégiant les citations plutôt que les exemples forgés de toute pièce, les rédacteurs choisissent de se mettre en retrait et font désormais porter « la responsabilité d'une éventuelle « légitimation » d'un emploi aux auteurs desdites citations. Cette prudence, saluée par Pöll (2002 : 79), semble être, pour l'heure, la seule attitude raisonnable pour la rédaction d'un dictionnaire différentiel.

Toutefois, pour dépasser cette prudence, nous cherchons à modéliser cette attitude afin d'obtenir des informations sur l'emploi du lexème et de les rendre explicites, sans pour autant les faire entrer dans une grille de marquage fixée. En effet, comme les différents systèmes de marquage existants ne répondent pas de manière entièrement satisfaisante à nos attentes, nous souhaitons réaliser une fiche signalétique du lexème qui soit utile pour l'utilisateur tout en cherchant à éviter une approche prescriptive.

Il s'avère donc nécessaire de définir les caractéristiques des régionalismes afin de présenter un état des lieux détaillé de leur usage par le plus grand nombre. Pour ce faire, nous avons décidé de limiter notre travail à une optique variationnelle « qualitative ». Cela signifie que nous ne nous livrons pas à des tests de fréquence dans différentes situations communicationnelles ; en effet, une procédure de ce type – par ailleurs fort intéressante – nécessiterait la mise en œuvre de moyens colossaux que nous ne sommes pas en mesure de déployer.

Pour cette raison, nous utilisons le corpus sélectionné uniquement dans l'optique d'une observation « systématique » des paramètres que nous jugeons pertinents. Notre corpus, comme le préconise Corbeil (1998 : 46), comportera

deux grandes sections : une première, constituée d'un « corpus de textes ou d'enregistrements pour l'observation et l'analyse de l'usage [du lexème] relevant de situations de communication variées », et un second ensemble composé de « commentaires sur la langue et le lexique, sorte de recueil de discours sur la langue ». La mise en place simultanée de ces deux volets doit permettre l'évaluation des lexèmes sur deux axes : leur norme de production témoignant de l'usage réel des locuteurs, et leur norme de réception indiquant quelle est l'évaluation du lexème par les locuteurs.

Nous obtenons ainsi plusieurs informations sur les traits distinctifs du régionalisme, notamment son caractère conscient ou inconscient, valorisé ou dévalorisé, intégré ou non intégré, connoté ou neutre. Pour ordonner ces traits, nous nous référons au travail de Hausmann (1989-1991) auquel nous empruntons les catégories de classement que nous transformons en grille d'évaluation.

Le travail s'organisera en un enchaînement de deux étapes qui se conditionneront mutuellement et qui permettront d'évaluer chaque lexème du corpus sélectionné selon dix paramètres variationnels qui reprennent le *Makromodell* proposé par Hausmann (1989-1991 : 649-657). La première étape nous permettra de déterminer les paramètres diachronique, diatopique et diafréquentiel, la deuxième les paramètres diatextuel, diamésique et diatechnique puis dianormatif et diaévaluatif³.

GRILLE D'ÉVALUATION

Notre corpus se composera d'une centaine d'occurrences sélectionnées parmi les entrées de la BDLPSuisse⁴ afin de constituer un échantillonnage représentatif de lexèmes établi selon des critères historiques, géolinguistiques et diaphasiques. Par le biais du traitement et de l'analyse de ce corpus restreint, nous chercherons à évaluer la possibilité de développer un *modèle* susceptible d'être ensuite éventuellement, et selon nos moyens, généralisé à l'ensemble de la nomenclature suisse romande.

DIACHRONIQUE, DIATOPIQUE, DIAFRÉQUENTIEL

Le premier stade de notre évaluation consiste à dresser le profil de chaque lexème, à savoir sa représentation géographique, sa fréquence et sa vitalité pour les différentes zones linguistiques de la Suisse romande.

En ce qui concerne le français régional de Suisse romande, certaines études ont déjà été menées, mais elles traitent uniquement de la place politique qu'occupe le français en Suisse romande et dans la francophonie (Knecht 1985; Lüdi *et al.* 1997; Chaudenson & Rakotomalala 2004) ainsi que des pratiques langagières et des représentations que les locuteurs ont de leur langue au niveau sociolinguistique (Manno 1993, 1994, 2002; Singy 1989, 1993, 1996, 2002, 2004). Par contre, aucune étude ou enquête n'a encore été réalisée pour

3 Pour des définitions détaillées de ces paramètres, v. Hausmann 1989-1991 : 649-657. Dans notre grille d'évaluation, nous laisserons de côté le paramètre diastratique; nous estimons en effet que le régionalisme est virtuellement disponible pour toutes les couches de la société régionale et que tout le monde est à même de l'utiliser lorsque le contexte le demande ou le favorise.

4 La BDLPSuisse contient l'ensemble des informations du DSR, lui-même réalisé à partir des matériaux provenant d'un large fonds documentaire déposé au Centre de dialectologie et d'étude du français régional de l'Université de Neuchâtel sous forme de fichier manuel (Schüle 1981). Ce fichier, régulièrement mis à jour, rassemble des attestations de la langue de la presse, de la littérature et de corpus oraux pour le français régional contemporain de la Suisse romande.

l'ensemble de la Suisse romande⁵ afin de déterminer l'influence de certains paramètres comme l'âge, la dichotomie ville-campagne, etc. (Knecht/Rubattel 1984). Une première étape sous forme d'enquête s'avère donc nécessaire et se fera sur le modèle de la démarche de Manno (1994), c'est-à-dire en combinant une « phase passive » et une « phase active » – la première portant sur la reconnaissance de lexèmes régionaux par les témoins, la seconde sur leur production spontanée – et en utilisant les paramètres sociolinguistiques habituels pour déterminer la vitalité des lexèmes sélectionnés⁶. Nous testerons leur vitalité diachronique, diatopique et diafréquentielle en les soumettant aux témoins représentatifs choisis en fonction de critères géolinguistique et diachronique.

DIATEXTUEL, DIAMÉSIQUE, DIAPHASIQUE, DIATECHNIQUE

La deuxième étape consiste à la fois à tester l'usage réel des locuteurs / scripteurs – la norme de production – et les représentations conscientes et inconscientes que le locuteur se fait de sa langue. Thibault (2000 : 78) souligne que :

un mot peut s'avérer parfaitement intégré dans un système linguistique du point de vue fonctionnel, sans nécessairement jouir d'une « légitimité » sociale ; car, besoin est-il de le rappeler, il peut y avoir un abîme entre la langue et la représentation que l'on s'en fait.

Pour évaluer cette production et cette réception, nous envisageons de mesurer les différentes dimensions dans lesquelles elles s'inscrivent – diatextuelle, diamésique, diatechnique et diaphasique – et qui n'entrent pas dans l'évaluation de la vitalité du lexème, et d'en déduire ses connotations stylistiques. Grâce au dépouillement de larges corpus textuels (Suistext⁷, archives de presse, Internet, etc.), nous sommes en mesure d'obtenir des informations sur le registre de langue du lexème (écrit, oral, formel, informel, officiel, etc.) ainsi que sur le contexte de production dans lequel il s'inscrit (du texte littéraire en passant par les archives de la presse régionale jusqu'aux registres les plus informels comme les forums de discussion Internet). Pour l'usage oral, nous intégrerons les informations récoltées lors de la « phase active » de l'enquête de terrain de la première étape.

DIAINTEGRATIF

Les indices nous permettant de mesurer la dimension intégrative d'un mot régional de nature « étrangère » (emprunts, germanismes, dialectalismes⁸) se si-

5 Toutefois, nous pouvons utiliser des enquêtes non publiées (comme celles menées par Schüle dans les années 1970 et par le Centre de dialectologie dans les années 1975-1981) et la documentation rassemblée pour la rédaction du DSR qui fournissent les caractéristiques géolinguistiques et diachroniques pour certains lexèmes.

6 Nous envisageons de constituer un réseau d'enquêtes englobant une vingtaine de localités couvrant l'ensemble de la Suisse romande urbaine et rurale, en tenant compte des structures géographique et socioculturelle de chaque canton. Des résultats statistiquement pertinents pourront être obtenus avec un ensemble de témoins comprenant environ 650 personnes (en dehors de Singy 1996, qui limite son étude au canton de Vaud, il n'existe pas d'enquête préalable de ce type). Comme chez Benoît & Michel (2001), nous superposerons aux critères du sexe et de l'âge ceux de l'appartenance socioprofessionnelle pour obtenir des informations significatives à ce sujet sans dépasser un nombre maîtrisable de témoins.

7 La base textuelle Suistext – constituée sur le même modèle que Frantext, Beltext et Québetext et élaborée sous l'impulsion du TVF – regroupe actuellement les œuvres complètes de 14 écrivains contemporains suisses romands.

8 Pour une étude sur l'origine des régionalismes lexicaux suisses romands, voir Thibault 1996.

tuent surtout au niveau de la graphie. Pour les germanismes, par exemple, la survivance de la majuscule initiale dans les corpus écrits témoigne d'un degré d'intégration moindre. À l'inverse, la présence d'emplois figurés ou d'une famille dérivationnelle étendue semble être autant de signes d'une bonne intégration.

DIANORMATIF

Nous testerons les connotations dianormatives grâce à l'analyse des productions suivantes : discours littéraires, discours normatifs, discours régionalistes, enquêtes sociolinguistiques et citations métalinguistiques. Ces différents aspects ont déjà été passés en revue par Thibault (1998).

Dans les textes littéraires d'auteurs romands, il semble désormais rare que l'on remette « en cause la légitimité des expressions régionales, mais la situation est très variée selon les auteurs et les ouvrages » (Thibault 1998 : 29)⁹.

Pour analyser le discours normatif, nous pensons d'une part faire appel aux cacologies – recueils de locutions ou de constructions considérées comme fautives – et aux chroniques de langue du XX^e siècle. D'autre part, nous procéderons au dépouillement des publications émanant des professionnels de l'écriture (journalistes et traducteurs), en particulier celles des organismes suivants : la section suisse de l'Union internationale de la presse francophone¹⁰ et le Fichier français de Berne (FFB)¹¹.

En ce qui concerne le discours régionaliste, il nous arrive par le biais de diverses émissions et œuvres qui font la part belle aux régionalismes (cf. bibliographie du DSR).

Pour l'approche sociolinguistique, nous chercherons, comme le suggère Thibault (1998 : 33), à « susciter un métadiscours chez le locuteur » et à récolter ainsi une information fiable sur son sentiment à l'égard de son lexique pour comprendre les rapports qu'il entretient avec sa langue et les conséquences identitaires que cela implique. Ces valeurs peuvent être obtenues par une vérification qualitative auprès des locuteurs par le biais d'entretiens semi-directifs (en adoptant une méthodologie qui s'inspire des démarches de Singy 1996, 2000 et de Pépin 1999) réalisés sur la base d'un échantillonnage de témoins sélectionnés à partir des résultats obtenus à l'étape précédente.

Les citations métalinguistiques sont également une source précieuse d'information. Elles consistent, comme le fait remarquer Thibault (1998 : 36-39), d'une part en des énoncés neutres dans lesquels l'accent est mis sur la différence entre un mot du français hexagonal et un mot du français de Suisse romande – sans toutefois pencher en faveur de l'un ou de l'autre – et d'autre part,

9 Comme nous l'a fait judicieusement remarquer Mme Thérèse Moreau, écrivaine suisse romande, il pourrait être profitable de savoir si les maisons d'édition ont des politiques définies à ce sujet.

10 Il s'agit d'une union professionnelle active anciennement nommée *Union internationale des journalistes et de la presse de langue française* qui publie le bulletin *Défense du français*. Pour des informations détaillées voir <http://www.presse-francophone.org/>.

11 Sur son site Internet (<http://www.fichier-français.ch/>), cet organisme se décrit de la manière suivante : « Le Fichier français est un cercle d'études ayant pour objet le bon usage de la langue française. Il lutte contre les altérations de la langue et fait front aux traductions hasardeuses ». Il publie régulièrement des «Fiches vertes» qui «ont pour vocation d'offrir un instrument pratique aux rédacteurs et aux traducteurs et de compléter ainsi les dictionnaires usuels». Il s'intéresse plus particulièrement aux germanismes et à la langue de l'administration fédérale.

en des énoncés hostiles ou, à l'inverse, favorables aux helvétismes¹². Toutefois, il serait restrictif de limiter les indications métalinguistiques aux seuls commentaires explicites; en effet, dans les corpus écrits, des marques telles que les guillemets ou les italiques peuvent également être considérées comme des signes de divergence.

DIAÉVALUTATIF

Pour le paramètre diaévaluatif, il n'est possible de rassembler que des informations éparpillées. Nous n'avons pas été en mesure de mettre sur pied un filtre permettant de saisir de manière satisfaisante la connotation d'un lexème. Il n'est donc pas possible, selon nous, d'atteindre pour l'instant des résultats aussi rigoureux que pour les autres catégories décrites ci-dessus. Toutefois, nous envisageons de faire appel aux séries synonymiques mises en contexte; il s'agit de mettre en œuvre le processus consistant à confronter le lexème régional à son équivalent en français standard. Cette méthode, qui a déjà servi pour évaluer le sociolecte d'un groupe donné (Lesigne 2000, Palazollo-Nöding 1987) devrait nous permettre d'obtenir des informations sur le degré de formalité du lexème en question et sur sa place dans la hiérarchie du champ synonymique. Pour effectuer cette vérification, et dans un but de cohérence des données, il semble profitable d'avoir recours aux témoins ayant participé à l'enquête de vitalité de la première étape.

1. EXEMPLES

Pour illustrer l'application concrète de cette grille d'évaluation, nous avons choisi un ensemble de lexèmes représentatifs de la diversité du lexique suisse romand. Il faut bien sûr insister sur le fait que, si la grille fait appel à des critères objectifs, l'analyse que nous en livrons n'est pas encore étayée par la démarche empirique. Nous avons, pour ces exemples, simplement inséré les informations partielles que nous possédons déjà tirées pour la majorité d'entre elles du DSR.

De ces fiches signalétiques se dégagent plusieurs cas de figure qui vont du classement classique du lexème selon un ou plusieurs paramètres variationnels, à un traitement particulier (marquage individuel des diverses acceptions du mot ou de certains éléments de sa famille dérivationnelle, modification du marquage, etc.).

1.1 CLASSEMENT SELON LES PARAMÈTRES VARIATIONNELS LES PLUS PRÉGNANTS

Tout d'abord, certains de ces lexèmes peuvent simplement être situés par rapport aux paramètres variationnels décrits plus haut :

a. *diachronique, diatopique, diafréquentiel*

régent, -e (n.m., f.) Instituteur, institutrice d'école primaire.

Vitalité : Suisse romande (SR) ; vieilli et rural.

Diatextuel, diamésique, diaphasique : encore très fréquent dans littérature contemporaine.

Dianormatif, diaévaluatif : marques typographiques; citations métalinguistiques.

¹² Il faut bien sûr, lors de l'analyse, garder à l'esprit les questions suivantes : qui tient ces discours et avec quelle intention ? Ces remarques concernent-elles tous les registres ?

Citations : [commentaire métalinguistique] L'« instit » chasse le vieux « *régent* » vaudois [titre]. (1982, *Tribune de Genève*, 16 février, p. 21) [presse, journaux, périodiques]

[commentaire métalinguistique] Dans les petits villages, l'instituteur est resté pour beaucoup le **régent** [en gras dans le texte]. Généralement, il s'occupe des grands. Les petits sont confiés à la *régente* [en ital. dans le texte]. Parfois, une *régente* de couture [en ital. dans le texte] complète l'équipe locale du corps enseignant. (1988, B. Chapuis, *Mots et sobriquets*. Terroir vivant, p. 34) [littérature]

Régent, -e est considéré comme vieilli et rural (on utilise plutôt de nos jours « maître, -esse ; instituteur, -trice »). Lorsqu'il est utilisé dans la littérature ou la presse contemporaines, c'est toujours en référence à une réalité passée ou désuète, bien souvent régionale ou rustique, et il est fréquent qu'il soit accompagné de commentaires métalinguistiques indiquant à la fois la distance diachronique et la régionalité du terme.

b. *diatextuel, diamésique, diaphasique, diatechnique*

roiller (v.intr.) Battre, frapper.

(imp.) Pleuvoir à verse.

Vitalité : SR.

Diatextuel, diamésique, diaphasique : fréquent à l'oral, mais très rare à l'écrit (littérature régionaliste).

Dianormatique, diaévaluatif : marques typographiques ; critiqué par quelques cacologies.

Citation : Les uns jouent d'un instrument : violon, piston, bugle, bombardon, grosse caisse, tambour, bigophone, etc., etc. Les autres marquent le rythme en *roillant* [en ital. dans le texte] sur de vieilles casseroles [...]. (1975, A. Itten, R. Bastian, *En ça... en là !*, p. 22) [littérature]

Attesté dans toute la Suisse romande, ce verbe est surtout présent à l'oral, mais très rarement à l'écrit. Critiqué par quelques cacologies, il apparaît cependant parfois dans la littérature régionaliste, le plus souvent marqué typographiquement.

c. *diaintégratif*

stamm (n.m.) Local de réunion d'une association, d'un parti politique, permanence.

Vitalité : SR.

Diatextuel, diamésique, diaphasique : très fréquent dans la presse.

Diaintégratif : emploi fréquent de la majuscule.

Dianormatif, diaévaluatif : marques typographiques ; critiqué par Défense du Français (germanisme).

Citation : M. Largiers m'entraînait chaque jour au « *Stamm* » du « Café de l'Ours ». « *Stamm* » veut dire en quelque sorte permanence [...]. C'est sans prétention, mais ces « *Stamm* » jouent un rôle non négligeable dans la vie des petites villes. (1965, R. Fell, *Dans l'été brûlant*, p. 9) [littérature]

Ce terme, attesté dans toute la Suisse romande, est très fréquent dans la presse. De manière générale, son origine germanique n'échappe à personne :

l'emploi répandu de la majuscule témoigne même d'une volonté de conserver au plus près le caractère allemand de l'emprunt.

Nous pouvons remarquer que la situation de ces lexèmes par rapport à des paramètres variationnels particuliers est bien souvent soulignée par la présence de marques dianormatives et/ou diaévaluatives : critiqués par les cacologies, commentaires métalinguistiques, guillemets, etc. C'est-à-dire que leur déviation ou leur connotation par rapport au standard culturel est sentie et mise en avant.

1.2 DIFFÉRENCE DE TRAITEMENT

Nous observons ensuite une différence de traitement entre les diverses acceptions d'un mot (sémantisme, emploi figuré, etc.) ou à l'intérieur d'une famille dérivationnelle. Il peut en effet arriver que dans une famille dérivationnelle un seul élément soit senti comme régional. Par ailleurs, un lexème peut parfois être marqué uniquement lorsqu'il apparaît dans un emploi figuré ou dans une expression; c'est notamment le cas d'*aguillage* et de sa famille dérivationnelle :

aguillage (n.m.) Amoncellement d'objets superposés, en équilibre plus ou moins instable. Assemblage d'objets fixés de façon peu soignée; réparation de fortune, rafistolage mal fait.

Vitalité : surtout cantons de Neuchâtel (NE), Vaud (VD) (deuxième sens semble également connu dans Jura (JU)).

Diatextuel, diamésique, diaphasique : littérature régionale (VD), presse et radio-télévision.

Diaintégratif : (dialectalisme) emplois figurés, grande famille dérivationnelle (déjà en grande partie attestée dans les dialectes).

Dianormatif, diaévaluatif : valorisé/critiqué par les cacologies; pour certains emplois figurés : commentaires métalinguistiques indiquant sa localisation (VD) et marques typographiques.

Citations : Ce genre de « contrat », passé implicitement dans la plupart des cas, produit des effets désastreux sur le peuple souverain, qui en flaire la teneur. Loin de s'en accommoder, il déteste ce genre d'« *aguillage* », sans pour autant toujours le sanctionner par un vote contraire. (1994, *Domaine Public*, 8 décembre, p. 45) [presse, journaux, périodiques]

[commentaire métalinguistique] Nous vivons une période fertile en miracles, du moins en apparence, car on voit de braves gens croire à des théories qui ne sont que de vulgaires « *aguillages* de noix sur un bâton », expression bien vaudoise qui rend parfaitement compte de la valeur de certaines théories où l'on fait davantage appel à la sentimentalité qu'à la raison. (1993, *Gazette de Lausanne*, 8 juillet) [presse, journaux, périodiques]

Ce régionalisme, surtout attesté à NE, VD, est accueilli principalement par la littérature régionale (essentiellement vaudoise), ainsi que par la presse et la radio-télévision. Cependant, la famille dérivationnelle *aguillage*, *aguiller*, *déguillage*, *déguillée*, *déguiller*, *raguillée*, etc. est alternativement valorisée et critiquée par les cacologies, ceci certainement pour les raisons énoncées par Humbert ([1852] 1983 : 12) : « ce verbe [aguiller], qui est d'un usage assez général, est très expressif et n'a point d'équivalent en français ». Cette absence d'équivalent dans le français de référence est un élément qui fausse la réception des lexèmes régionaux, puisque ceux-ci semblent de fait bénéficier d'une légitimité

« naturelle » en venant combler un manque. De plus, au niveau de la production, seuls les emplois figurés ou plaisants de cette importante famille dérivationnelle sont marqués typographiquement et/ou identifiés comme vaudois.

Pour un autre lexème, le verbe *encoubler*, il existe une différence entre le traitement des emplois transitif et pronominal :

encoubler (v.tr.) Déranger, importuner, gêner, embarrasser.

(v.pron.) Trébucher, s'accrocher les pieds dans qqch., s'empêtrer.

Vitalité : SR.

Diatextuel, diamésique, diaphasique : littérature, presse, enquêtes.

Diainfégratif : (dialectalismes) nombreux emplois figurés.

Dianormatif, diaévaluatif : emploi pronominal : critiqué / valorisé par les cacologies. Emploi transitif : absent des cacologies ; marques typographiques.

Citations : On ne peut pas circuler, on doit aller à l'écurie, à la cuisine, vers les cochons, alors ces gens, ils nous « *encoubent* » comme on dit. (1975, P. Hugger, *Le Jura Vaudois*, p. 227) [littérature]

Pour que les internautes juniors ne *s'encoubent* pas virtuellement dans la grande toile, ils ont intérêt à commencer de surfer accompagnés. (1997, *Construire*, 27 août, p. 13) [presse, journaux, périodiques]

[commentaire métalinguistique] Sans parler de ceux qui veulent parler si vite qu'ils *s'encoubent* (oui, je sais, à la télé on dit : « butent ») sur leurs mots... Beau gâchis ! Celui qui a honte de ses origines est tout simplement un pauvre type. (1997, *Construire*, 12 novembre, p. 20) [presse, journaux, périodiques]

Au niveau de la vitalité tout d'abord, l'emploi transitif semble en voie de disparition, tandis que l'emploi pronominal est fréquent et actuel. Cette différence de vitalité est soulignée par l'utilisation de la forme *s'encoubler* dans des énoncés concernant des thèmes modernes comme Internet (cf. citation) ainsi que par l'absence de mention de l'emploi transitif dans les cacologies lorsqu'elles critiquent ce verbe.

Cette tendance est confirmée par le traitement observable dans les attestations que nous possédons, à savoir l'apparition fréquente de marques typographiques pour l'emploi transitif et leur absence pour la forme pronominale.

Ainsi, *s'encoubler* semble être un régionalisme inconscient, ce que confirment les quelques témoignages ponctuels que nous possédons déjà de locuteurs. Et lorsqu'il est repéré, le discours puriste et le discours régionaliste tombent d'accord pour le valoriser, comme le fait Dudan (1945), qui le juge « expressif et excellent » (voir à ce propos la citation métalinguistique ci-dessus).

1.3 MODIFICATIONS DU MARQUAGE

Et enfin, nous remarquons des modifications du marquage lorsque des mots changent de domaine d'emploi et/ou de statut (diatechnique, diatopique, dianormatif). C'est notamment le cas d'un autre régionalisme inconscient : le verbe *prétérir* :

prétérir (v.tr.) Léser, causer du tort à, porter préjudice à.

Vitalité: SR.

Diatextuel, diamésique, diaphasique : beaucoup plus fréquent dans presse, écrits administratifs ou politiques que dans littérature.

Dianormatif, diaévaluatif : proscrit puis réhabilité par FFB.

Citations : Nous espérons que le Conseil d'Etat pourra infirmer nos craintes, faute de quoi notre agriculture neuchâteloise serait *prétéritée* par rapport à celle d'autres cantons. (1972, *Bulletin officiel des délibérations du Grand Conseil*, Neuchâtel, 11 décembre, p. 1050) [textes administratifs ou officiels]

En principe *prétérités* par rapport aux télécopieurs, qui traitent directement le papier, ces engins s'avèrent irrésistibles pour les pécéistes [utilisateurs de micro-ordinateurs] précisément parce qu'ils permettent de faxer sans papier ! (1993, *Tribune de Genève*, 10 mars, p. 21) [presse, journaux, périodiques]

Cet emploi est légitimé par une presse de qualité et par la langue administrative. De plus, comme le fait remarquer le DSR, ce verbe a été réhabilité comme « ayant sa place dans le vocabulaire français » depuis une vingtaine d'années par FFB qui l'avait proscrit auparavant. Il est donc très intégré dans un milieu formel et officiel. Diffusé par l'intermédiaire de la langue des notaires de Suisse romande, ce terme alors appliqué uniquement à des personnes ou des institutions, tend à s'ouvrir à des emplois moins figés, moins formels, et s'utilise désormais aussi pour toute réalité dans le sens plus général de « désavantager ». L'emploi de ce terme semble donc s'étendre à un domaine plus large et perdre son sens technique.

Huitante bénéficie d'un statut un peu comparable à celui de *prétériter* par le fait qu'il est la forme communément usitée, pour certains cantons uniquement (VD ; VS ; FR), dans l'usage oral, scolaire et administratif :

huitante (adj. num. card. inv.) Quatre-vingts.

Vitalité : VD, VS, Fribourg (FR) (actuellement usité dans la téléphonie mobile romande).

Diatextuel, diamésique, diaphasique : presse, littérature, écrits administratifs.

Dianormatif, diaévaluatif : toléré / valorisé / critiqué ; commentaires métalinguistiques (souvent valorisé).

Citations : À Zürich, jusqu'au 31 juillet, les sculptures de Duane Hanson (1925-1996) dialoguent avec des peintures d'artistes américains des années soixante aux années *huitante* issues des collections du Kunsthaus. (2003, *Domaine Public*, 4 juillet, p. 8) [presse, journaux, périodiques]

Modifier l'article 6 des statuts afin qu'il ait la teneur suivante : « Le capital-actions de cinq cent trente millions cinq cent mille francs (CHF 530'505'000.-) est divisé en huit millions quatre cent *huitante* huit mille et *huitante* (8'488'080) actions de soixante-deux francs et cinquante centimes (CHF 62.50) nominal chacune, nominatives, entièrement libérées. » (Ordre du jour de l'assemblée générale extraordinaire des actionnaires de la banque cantonale vaudoise, février 2003, site Internet www.bcv.ch) [écrits administratifs]

On peut observer dans la production écrite que *huitante* est toujours utilisé sans signe diacritique, mais que l'équivalent de référence *quatre-vingt* domine très largement (même chez les auteurs originaires VD ; VS ; FR). Les discours normatifs sont partagés à son propos : les auteurs tour à tour le valorisent (« de la meilleure veine latine », Dudan 1945), le tolèrent ou le critiquent. Les arguments sont divers et appellent à l'intercompréhension ou à la tradition, comme le FFB qui le tolère mais conseille qu'on use de l'équivalent *quatre-vingts* « dans

les contacts avec les Français ». Quant à la *Défense du français*, elle le considère comme un archaïsme et suggère d'employer plutôt *quatre-vingts* « un des derniers souvenirs linguistiques de nos ancêtres les Gaulois, qui comptaient par vingtaines ». La question semble donc être sans ambiguïté : *huitante* est de « chez nous », *quatre-vingts* de « chez eux », et ces termes ne sont pas en concurrence. Seule la localisation les distingue. Le fait que cet emploi ne soit pas stigmatisé dans les régions où son usage est signalé a permis à ce lexème d'entrer dans l'usage de certaines sociétés (téléphonie mobile, Banque Cantonale Vaudoise, etc.) et semble favoriser sa propagation à toute la Suisse romande.

2. RÉSULTATS ATTENDUS

Comme nous l'avons vu, pour certains lexèmes suisses romands, il existe, comme le dit Thibault (1998 : 40), une « légitimité qui varie [...] selon l'époque et le type de discours, la nature des particularités en cause (accents, mots, expressions), l'énonciateur et le contexte énonciatif ».

De fait, nous espérons parvenir à démontrer qu'il existe une caractérisation et une stratification des lexèmes qui n'est pas uniquement fonction de leur rapport au français de référence et qui peut varier dans l'espace. Cela devrait nous permettre, non seulement de cerner l'existence de normes locales et suprarégionales en français régional de Suisse romande – sans que cela implique, cependant, « l'existence d'un français national suisse romand dans la conscience des locuteurs »¹³ (Thibault 1998 : 25; Gleßgen & Thibault 2005 : vj) – mais aussi de mettre en lumière l'existence de centres d'irradiation qui jusqu'à présent n'étaient que pressentis à l'échelle régionale (Müller 1975)¹⁴.

Les résultats obtenus nous permettront alors de vérifier l'hypothèse traditionnelle selon laquelle le prestige de la variété régionale change selon les régions – à l'intérieur desquelles la dichotomie ville-campagne est également susceptible de se révéler comme pertinente – et qu'il existe un développement d'espaces référentiels à l'échelle cantonale (Singy 1996). Il s'agira en particulier d'examiner dans quelle mesure la Suisse romande reste polycentrique par rapport à ses normes linguistiques, ou, à l'inverse, s'il est possible de déceler le début de l'émergence d'un espace culturel romand commun qui partage les mêmes références sous l'influence des médias modernes.

Alors que, jusqu'à présent, la réflexion sur les marques d'usage prenait toujours comme point de départ et comme objet d'étude la lexicographie du français de référence, nous pensons qu'une approche théorique de la notion de marque d'usage appliquée à la lexicographie différentielle pourra nous permettre de développer un cadre de référence pour mieux appréhender la nature intrinsèque du français régional de Suisse romande et l'outil théorique de sa description. Ceci devrait nous permettre de dépasser l'idée traditionnelle véhiculée par la notion de « régionalismes de bon aloi » pour considérer que ces lexèmes s'inscrivent dans un diasystème complet et non simplement en marge de la langue.

13 En effet, alors qu'il existe actuellement au Québec une variété légitime considérée comme « le français de référence québécois », il n'y a jamais eu la moindre tentative de définir et de standardiser une variété de « français national » pour la Suisse romande qui, linguistiquement, s'est toujours considérée comme une sorte de prolongement naturel du français de France (v. Schüle 1981).

14 Knecht & Rubattel (1984 : 144) écrivent à ce sujet que « Müller (1975 : 228) s'avance beaucoup en prétendant que Genève, Lausanne et Fribourg sont de tels centres d'irradiation. C'est possible, mais il n'en existe aucune preuve ».

RÉFÉRENCES BIBLIOGRAPHIQUES

- BENOÎT Michèle et MICHEL Claude, 2001, *Le Parler de Metz et du Pays Messin*, Metz, Éditions Serpenoise.
- CHAUDENSON Robert et RAKOTOMALALA Dorothée (éds), 2004, *Situations linguistiques de la francophonie. État des lieux*, Paris, AUF.
- CORBEIL Jean-Claude, 1998, « Les marques d'usage comme technique de description des aspects connotatifs du lexique », dans L. Mercier et C. Verreault (éds), *Les Marques lexicographiques en contexte québécois. Actes de la Table ronde tenue à Montréal les 3 et 4 novembre 1994*, Québec, Gouvernement du Québec, p. 29-50.
- DUDAN Camille, 1945, *Le français, notre langue, Quatrième série*, Bienne, Éditions du Chandelier.
- GLESSGEN Martin-Dietrich et THIBAUT André (éds), 2005, *La Lexicographie différentielle du français et le « Dictionnaire des régionalismes de France » : Actes du Colloque en l'honneur de Pierre Rézeau*, Strasbourg, Presses Universitaires de Strasbourg.
- HAUSMANN Franz Joseph et alii (Hrsg.), 1989-1991, *Wörterbücher: Ein internationales Handbuch zur Lexicographie / Dictionaries: An international encyclopædia of lexicography / Dictionnaires : Encyclopédie internationale de lexicographie*, Berlin u. New York, Walter de Gruyter.
- HUMBERT Jean, [1852] 1983, *Nouveau Glossaire genevois*, Genève, Slatkine Reprints.
- KNECHT Pierre, 1985, « La Suisse Romande », dans R. Schläpfer (éd.), *La Suisse aux quatre langues*, Genève, Éditions Zoé, p. 125-169.
- KNECHT Pierre et RUBATTEL Christian, 1984, « À propos de la dimension sociolinguistique du français en Suisse romande », dans G. Antoine et J. Chaurand (éds), *Le français moderne*, 52 (3/4), p. 138-150.
- LESIGNE Hubert, 2000, *Les Banlieues, les profs et les mots : essai de lexicologie sociale*, Paris, L'Harmattan.
- LÜDI Georges et al., 1997, *Le Paysage linguistique en Suisse. Recensement fédéral de la population 1990*, Berne, Office fédéral de la statistique.
- MANNO Giuseppe, 1993, « Français argotique et français régional en Suisse romande », dans G. Hilty (éd.), *Actes du XX^e Congrès International de Linguistique et philologie romanes, Université de Zürich (6-11 avril 1992), Tome III, Section V*, Tübingen et Bâle, Francke Verlag, p. 545-556.
- MANNO Giuseppe, 1994, *Le Français non conventionnel en Suisse romande*, Berne, etc., Peter Lang.
- MANNO Giuseppe, 2002, « La dynamique interne propre au français régional de Suisse romande : réflexions théoriques et méthodologiques autour d'un facteur sous-estimé », dans P. Singy (éd.), *Le Français parlé dans le domaine francoprovençal : une réalité plurinationale, Sciences pour la communication*, 66, Berne, etc., Peter Lang, p. 83-112.
- MÜLLER Bodo, 1975, *Das französische Gegenwart*, Heidelberg, Carl Winter.
- PALAZZOLO-NÖDING Brigitte, 1987, *Drei Substandardregister im Französischen : familier, populaire, vulgaire*, Frankfurt am Main, Haag und Herchen.

- PÉPIN Nicolas, 1999, *Entre identité et altérité. Quelques représentations linguistiques d'une famille de Suisse romande*, Mémoire de licence, Faculté des Lettres, Neuchâtel, Université de Neuchâtel.
- POIRIER Claude, 2005, « La dynamique du français à travers l'espace francophone à la lumière de la Base de données lexicographiques panfrancophone », *Revue de linguistique romane*, 69, p. 483-51.
- PÖLL Bernhard, 2002, « Le français en Suisse romande : à propos de conceptualisations profanes et scientifiques du fait régional », dans P. Singy (éd.), *Le français parlé dans le domaine francoprovençal : une réalité plurinationale*, Berne, etc., Peter Lang, p. 67-82.
- QUÉMADA Bernard, 1990, « Trésor informatisé des vocabulaires francophones », dans A. Clas et B. Ouoba (éds), *Visages du français. Variétés lexicales de l'espace francophone*, Paris, J. Libbey.
- REY Alain, 1983, « Normes et dictionnaires », dans E. Bédard et J. Maurais (éds), *La Norme linguistique*, Paris, Le Robert.
- SCHÜLE Ernest, 1981, « Le français régional de Suisse », *Actes du colloque Les français régionaux, Québec 1979*, Conseil international de la langue française, p. 181-193 et 232-238.
- SINGY Pascal, 1989, « Français régional et fonction signum social », *Bulletin de la section de linguistique de la faculté des Lettres de Lausanne*, 10, p. 17-102.
- SINGY Pascal, 1993, « Français régional et individuation linguistique : latitudes et limites », dans G. Hilty (éd.), *Actes du XX^e Congrès International de Linguistique et philologie romanes, Université de Zürich (6-11 avril 1992), Tome III - Section V*, Tübingen et Bâle, Francke Verlag, p. 616-625.
- SINGY Pascal 1996, *L'Image du français en Suisse romande. Une enquête sociolinguistique en Pays de Vaud*, Paris, L'Harmattan.
- SINGY Pascal (éd.), 2002, *Le Français parlé dans le domaine francoprovençal : une réalité plurinationale*, Sciences pour la communication, 66, Berne, etc., Peter Lang.
- SINGY Pascal (éd.), 2004, *Identités de genre, identités de classe et insécurité linguistique*, Berne, etc., Peter Lang.
- THIBAUT André, 1996, « Québécoismes et helvétismes : éclairages réciproques », dans T. Lavoie (éd.), *Français du Canada - français de France. Actes du 4^e Colloque International de Chicoutimi, Québec, du 21 au 24 septembre 1994*, Tübingen, Niemeyer, p. 333-376.
- THIBAUT André, 1998, « Légitimité linguistique des français nationaux hors de France : le français de Suisse romande », *Revue québécoise de linguistique*, 26 (2), p. 25-42.
- THIBAUT André, 2000, « Le traitement des emprunts dans le DSR : aperçus théoriques et méthodologiques », dans D. Latin et C. Poirier (éds), *Contacts de langue et identités culturelles. Actes des 4^{es} journées scientifiques du réseau « Études du français en Francophonie »*, Québec, Les Presses de l'Université Laval, p. 69-84.
- THIBAUT André et KNECHT Pierre, [1997] 2004, *Dictionnaire suisse romand. Particularités lexicales du français contemporain*, Genève, Zoé.

MISE EN RELATION DES PARTICULARISMES QUÉBÉCOIS
ET HEXAGONAUX DANS LES DICTIONNAIRES USUELS
DU FRANÇAIS :
VOCABULAIRE POLITIQUE ET LIMITES
DU MARQUAGE TOPOLECTAL

par Mireille ELCHACAR, Université de Sherbrooke / Université de Cergy-Pontoise¹
et Louis MERCIER, Université de Sherbrooke

Résumé — L'histoire de la lexicographie québécoise est aussi l'histoire d'une affirmation identitaire grandissante. Après s'être limitée à des ouvrages différentiels ou encore normatifs, la lexicographie québécoise semble être mûre pour un ouvrage général. Ce type d'ouvrage implique la cohabitation d'usages québécois et français. Les marques topolectales peuvent dans cette optique être un bon outil pour faciliter la cohabitation des deux variétés de français. Le corpus à l'étude est le vocabulaire politique, champ lexical hautement identitaire. L'étude ici présentée porte sur la manière dont les marques topolectales peuvent être utilisées pour mettre en relation les vocabulaires politiques québécois et français dans un dictionnaire général. Les marques topolectales s'appliquent-elles à toutes les situations de variation? Nous nous pencherons sur cette question en puisant des exemples dans le vocabulaire politique.

Mots clés — français canadien, projet FRANQUS, lexicographie québécoise, norme régionale, marque topolectale.

Abstract — Quebec lexicography has traditionally been limited to two different types of work: those using a differential approach (focussing only on Quebec French), and those using a normative approach (focussing on the differences between Quebec French and French from France). We are now at a stage where it is possible to have a dictionary that encompasses both French from France and Quebec- French varieties: this is the aim of the FRANQUS's project. Labels referring to places can be a very good tool to facilitate this co- existence within the same volume. As an example the authors show how political vocabulary, which defines highly different lexical fields in France and in Quebec, can actually be presented in a general dictionary encompassing both varieties.

¹ Les auteurs sont membres du Centre d'analyse et de traitement informatique du français québécois (CATIFQ) et collaborent au projet de dictionnaire général en cours de réalisation à Sherbrooke. Mireille Elchacar remercie de leur appui financier le Conseil de recherches en sciences humaines du Canada, la Faculté des lettres et sciences humaines de l'Université de Sherbrooke ainsi que le CATIFQ.

Keywords — Canadian French, Quebec lexicography, FRANQUS project, regional norms, labels referring to places.

INTRODUCTION

La lexicographie québécoise s'efforce depuis un demi-siècle déjà à concevoir un dictionnaire général du français qui réponde adéquatement aux besoins langagiers des francophones du Québec et du Canada, un ouvrage qui respecte leur vision normative, qui prenne en compte leur expérience et leur point de vue de Nord-Américains, et auquel, pour toutes ces raisons, ils puissent s'identifier. Malgré leur intérêt manifeste pour les *Petit Robert* (NPR) et *Petit Larousse* (PL), les Québécois se rendent bien compte que ces ouvrages n'ont pas été conçus pour eux.

Le défi est important puisque le contexte linguistique québécois est passablement différent du contexte linguistique français. Si, en raison de l'association historique qui existe entre la langue française, la France et la culture française, les lexicographes de l'Hexagone peuvent décrire le français en se basant essentiellement ou presque essentiellement sur son usage en France, la tâche est nettement plus complexe pour les lexicographes québécois. Elle l'est notamment en raison de la problématique normative et surtout en raison de la difficulté qui subsiste encore aujourd'hui à obtenir un consensus social sur la mise en relation des normes endogène (québécoise) et exogène (véhiculée par les ouvrages produits en France). Devant la difficulté de la situation, les lexicographes québécois sont forcés d'innover, et c'est ainsi qu'ils contribuent à l'avancement de la lexicographie française ou, plus justement, de la lexicographie francophone.

Une équipe d'universitaires québécois, basée principalement à l'Université de Sherbrooke, travaille actuellement à la rédaction d'un dictionnaire du français standard en usage au Québec (il s'agit du projet *FRANQUS*). Prévu pour 2011, ce dictionnaire général québécois sera le quatrième à paraître depuis le milieu du XX^e siècle. Comme on le verra plus loin, ce nouvel ouvrage se démarque à la fois des dictionnaires québécois précédents et des dictionnaires usuels du français de facture hexagonale (NPR et PL) par sa façon d'aborder les cas de variation topolectale observables entre l'usage québécois et l'usage hexagonal.

Cet article comporte deux parties complémentaires. La première partie dresse un bref historique de la lexicographie généraliste québécoise en matière de marquage et propose une caractérisation des diverses approches exploitées. Dans la seconde partie, nous nous intéressons à quelques emplois caractéristiques du vocabulaire politique québécois pour illustrer les limites du marquage topolectal.

1. LEXICOGRAPHIE QUÉBÉCOISE ET MARQUAGE TOPOLECTAL : BREF HISTORIQUE ET PERSPECTIVE ACTUELLE

1.1 DANS LE *BÉLISLE* (1957, 1971 ET 1979)

Le premier dictionnaire général produit au Québec est le *Dictionnaire général de la langue française au Canada* de Louis-Alexandre Bélisle. Cet ouvrage, dont la première édition date de 1957, est une adaptation nord-américaine du Littré-Beaujean².

2 Le *Littré-Beaujean* ou *Petit Littré* est un abrégé du dictionnaire de Littré.

SUCRE (lat. *saccharum*) *n.m.* Suc très doux que l'on tire principalement de la canne à sucre, de l'eau d'érable et de la betterave, et que l'on transforme, au moyen du feu, en une substance cristallisée soluble dans l'eau. ✦ Fig. *Le sucre des paroles*, paroles flatteuses, doucereuses. ✦ Fig. *Un apothicaire sans sucre*, un homme qui manque des objets nécessaires à sa profession. ✦ Fig. *C'est tout miel et tout sucre*, se dit d'une personne doucereuse. ✦ Popul. *C'est un sucre*, en parlant de fruits très doux. ✦ *Sucre brut*, sucre qui, ayant été cuit, n'est pas encore raffiné. ✦ *Sucre raffiné*, sucre brut qui a été blanchi par le raffinage. ✦ *Sucre en pain*, masse de sucre raffiné, à laquelle on a donné une forme conique. ✦ Famil. *En pain de sucre*, en forme de cône: *il a la tête en pain de sucre*. ✦ *Sucre candi*, voy. *candi*. ✦ *Sucre cristallisable*, sucre de canne, de betterave, etc. ✦ *Sucre liquide ou incristallisable*, sucre de fruit. ✦ *Fruits confits à plein sucre*, fruits confits une livre avec une livre de sucre. ✦ *Confitures à mi-sucre*, confitures faites avec la moitié du sucre qu'on y met d'ordinaire. ✦ *Sucre d'orge*, voy. *orge*. ✦ *Sucre de pomme*, espèce de sucre candi fait en y mêlant du jus de pomme. ✦ *Sucre du foie*, *sucre de diabète*, *sucre urinaire*, principe sucré qui existe à l'état normal dans le parenchyme du foie, et dans l'urine en certains états pathologiques. ✦ En chim. Tout corps qui peut être transformé en alcool. ✦ *Sucre de lait* principe qui existe dans le lait de tous les mammifères. ✦ *Sucre blanc*, sucre raffiné. ✦ *Sucre brun*, cassonade, sucre roux. ✦ *Sucre du pays*, sucre d'érable. ✦ *Sucre de sève*, sucre d'érable plastique et plus foncé que l'on obtient vers la fin de la coulée des érables. ✦ *Les sucres*, *le temps des sucres*, saison, période du printemps, pendant laquelle on fabrique le sucre d'érable. ✦ *Un pain de sucre d'érable* est un gros cube de ce sucre. ✦ *Aller aux sucres*, aller se régaler de sucre, de tire d'érable à la cabane à sucre. ✦ *Travailler aux sucres*, travailler à la fabrication du sucre, du sirop, de la tire d'érable. ✦ *Partie de sucre*, partie de plaisir consistant à *aller aux sucres*. ✦ *Sucre à la crème*, bonbon fait d'un mélange de sucre et de crème.

Reprise
intégrale du
Littré-Beaujean
(ou *Petit Littré*)



Bloc d'emplois
québécois
ajoutés par
Bélisle

Tableau 1. L'article sucre du Bélisle (1957)

Comme on peut le voir à partir de l'article *sucre*, Bélisle reprend d'abord intégralement le contenu de l'article du dictionnaire français avant d'amorcer la description des particularismes canadiens-français. Les canadianismes ne sont pas intégrés logiquement à la description française de base, mais renvoyés à la fin, ce qui a pour effet de les marginaliser : par exemple, le canadianisme *sucre blanc* (2^e section) aurait pu être associé à *sucre raffiné* (1^{re} section), mais la frontière est étanche entre les parties « française » et « canadienne ». La deuxième section est annoncée par un symbole jouant le rôle d'une marque toponymique pour tout le bloc d'emplois qui suit, du moins dans l'édition de 1957. À partir de la deuxième édition de 1971, Bélisle adopte la classification normative en trois catégories proposée par l'Office de la langue française, qu'il identifie par trois symboles distincts (Tableau 2).

Malgré son ouverture plus grande aux usages canadiens, l'approche descriptive adoptée par Bélisle en ce qui a trait à la variation toponymique ressemble à celle qui caractérise l'approche actuelle des dictionnaires usuels de facture hexagonale. Dans le premier dictionnaire général canadien, comme c'est le cas dans le *NPR* et le *PL*, le français comme langue et la variété hexagonale de français (incluant ses particularismes = FF) se confondent et servent de base à la description; les canadianismes (ou québécoismes selon la terminologie actuelle = FQ) sont identifiés comme des écarts par rapport à cette base (schéma 1).

La période de transition dont je parlais dans ma première édition se caractérise maintenant par une politique du juste milieu. L'Office de la Langue française du Québec a établi les normes qui doivent servir à nous guider dans l'emploi des mots de notre cru. Avec sa collaboration, j'ai établi dans cette édition trois catégories de canadianismes.

1° Les CANADIANISMES DE BON ALOI, qui représentent des réalités nord-américaines dignes d'être reconnues sur le plan international (faune, flore, toponymie, poids et mesures, droit, etc.) et qui sont identifiés par le signe ©

2° Les CANADIANISMES POPULAIRES ET FOLKLORIQUES, représentés dans notre langue familière par certains archaïsmes et autres mots ou expressions qui ne sont pas reçus dans le bon usage de la francophonie. Les mots de cette catégorie sont identifiés dans ce dictionnaire par le signe ♣

3° Les ANGLICISMES, BARBARISMES et autres impropriétés pour lesquels existent dans le français international de réels équivalents qui sont donnés en guise de définitions. Ces mots qu'il serait généralement souhaitable de faire disparaître de la langue sont identifiés ici par le signe ⊗

Ce classement répond aux normes établies par l'Office de la langue française et aux jugements de valeur portés par sa Commission consultative sur les canadianismes que contient la présente édition. Il ne faudrait toutefois pas voir dans les distinctions établies par ces signes un jugement irrévocable – pas plus d'ailleurs que ne l'était ma première classification. En fait, plusieurs centaines de mots et d'expressions portant la fleur de lis qui les classait comme CANADIANISMES ou NÉOLOGISMES en 1957 se retrouvent aujourd'hui dans les plus récents dictionnaires usuels de la francophonie.

Tableau 2. Extrait de la préface de la deuxième édition du Bélisle (1971)



Schéma 1 : Approche descriptive et marquage topolectal dans le Bélisle

1.2 DANS LE *DFP* (1988) ET LE *DQA* (1992 ET 1993)

Un deuxième dictionnaire général produit au Québec paraît en 1988. Il s'agit du *Dictionnaire du français Plus à l'intention des francophones d'Amérique* (*DFP*). L'ouvrage résulte de l'adaptation d'une matrice hexagonale de la maison Hachette, réalisée notamment par Claude Poirier et deux collaborateurs du *Trésor de la langue française au Québec*. Avec la parution du *DFP*, on assiste à un changement radical en ce qui a trait au marquage topolectal. Comme l'illustre le schéma suivant, l'approche est inversée : dans la mesure où ce dictionnaire choisit d'axer sa description sur le français en usage au Québec (abordé comme un système fonctionnel), c'est maintenant l'usage québécois du français qui se confond avec la langue française pour servir de base à la description. Et ce sont les francismes, c'est-à-dire les particularismes de l'usage hexagonal, qui sont présentés comme des éléments externes et qui font l'objet d'un marquage topolectal : on les fait précéder de la parenthèse (France).



Schéma 2 : Approche descriptive et marquage topolectal dans le *DFP* et le *DQA*

Si l'on compare l'article *sucre* du *DFP* (tableau 3 ci-dessous) à celui du *Bélisle*, on constate que les emplois marqués comme canadiens dans le *Bélisle* (*sucre blanc*, *sucre brun*, *sucre du pays*, *sucre à la crème*, etc.) sont logiquement intégrés à l'article sans aucune marque identificatrice. En revanche, deux emplois sont identifiés comme caractéristiques du français hexagonal (*sucre glace* et *sucre semoule*).

En 1992, la maison DicoRobert lance sur le marché québécois le *Dictionnaire québécois d'aujourd'hui* (*DQA*). Réalisé sous la direction de Jean-Claude Boulanger, le *DQA* résulte lui aussi de l'adaptation d'un dictionnaire général produit en France (le *Micro-Robert*). Comme on peut le voir en comparant l'article *sucre* du *DQA* (v. ci-dessous) avec l'article correspondant du *DFP*, ce troisième dictionnaire général québécois maintient l'approche initiée par son prédécesseur : le seul emploi marqué sur le plan topolectal est le francisme *sucre glace*.

Le *DFP* (qui n'a pas connu de réédition) et le *DQA* (qui n'a pas été réédité après 1993) ont abouti à un échec commercial alors que, depuis sa première édition en 1988, le principal dictionnaire de difficultés québécoises, le *MultiDictionnaire* ou *Multi* jouit d'un succès commercial constant. À l'instar du *Bélisle*, le *Multi* fait précéder les québécismes d'un symbole de fleur de lis et n'identifie

DFP 1988	DQA 1992
<p>I.1. Substance alimentaire de saveur douce que l'on tire principalement de la betterave et de la canne à sucre. <i>Sucre raffiné. Sucre blanc, sucre brun. Sucre en poudre (France) Sucre glace en poudre très fine. (France) Sucre semoule, en poudre grossière. Carré, morceau de sucre. — Vin de sucre, obtenu par la fermentation de marcs additionnés d'eau sucrée. ▷ Fam. Morceau de sucre. ▷ Loc. fig. Casser du sucre sur le dos de qqn : dire du mal de lui. — Être tout sucre et tout miel, très doucereux. — Fam. Cet enfant n'est pas en sucre !, il n'est pas si délicat, si fragile.</i></p>	<p>I. 1. Substance alimentaire, blanche, cristallisée (saccharose), de saveur douce, soluble dans l'eau. => gluc(o)-, sacchar-. <i>Sucre de canne, de betterave. Sucre en morceaux, cristallisé, en poudre ou à glacer. — (France) Sucre glace, finement broyé. Anglic. Sucre brun, roux. => cassonade. — Régime sans sucre des diabétiques. Un sucre, un carré de sucre, un morceau de sucre. Pince à sucre. — Loc. Être TOUT SUCRE TOUT MIEL: se faire très doux, doucereux. CASSER DU SUCRE sur le dos de qqn : en dire du mal en son absence.</i></p>
<p>2. <u>Sucre à la crème : confiserie fondante traditionnelle à base de sucre et de crème, qu'on sert découpée en carrés. Un plateau, un morceau de sucre à la crème.</u> — <i>Sucre d'orge : sucre aromatisé roulé en bâton.</i></p>	<p>2. Cette substance, préparée en confiserie. — SUCRE À LA CRÈME : friandise fondante faite de sirop, de sucre ou de cassonade qu'on fait bouillir avec de la crème puis refroidir. Des carrés de sucre à la crème. <i>SUCRE D'ORGE : sucre cuit et parfumé, présenté en petits bâtons (confiserie).</i></p>
<p>3. CHIM Glucide*.</p>	<p>3. Chimie. Corps ayant une constitution voisine du saccharose (ex. <i>glucose</i>).</p>
<p>II. (Dans le voc. de l'exploitation de l'érable à sucre.)</p> <p>1. <u>Sucre (d'érable) (vx sucre du pays). Sucre de sève. Sucre mou.</u></p> <p>2. <u>Loc. Cabane à sucre.</u> <i>Faire les sucres. Le temps, la saison des sucres (ou ellipt. les sucres). Aller aux sucres. Partie de sucre.</i></p>	<p>II. Dans l'exploitation de l'érable à sucre.</p> <p>1. <u>Sucre (d'érable ou vx du pays). Sucre mou, sucre dur.</u></p> <p>2. <u>Cabane* (à sucre).</u></p> <p>3. <u>Au plur. LES SUCRES.</u> — <i>Loc. La saison, le temps des sucres ou ellipt. les sucres. — Faire les sucres. — Aller aux sucres. — Partie de sucre.</i></p>

Tableau 3. Les articles sucre du DFP et du DQA³

³ Les emplois caractéristiques du français québécois sont mis en gras et soulignés, ceux du français hexagonal sont en grisé. Tous les emplois de la section II relevant du français québécois, nous n'en reproduisons que les sous-entrées.

pas les francismes⁴. L'échec du *DFP* et du *DQA* a clairement démontré que leur approche lexicographique, pourtant bien fondée linguistiquement, ne répondait pas adéquatement aux attentes actuelles de la société québécoise, ou du moins aux attentes des milieux québécois les plus préoccupés par la question normative :

l'accueil enthousiaste qu'a reçu [le *DFP*] s'est tempéré d'inquiétude quand on s'est rendu compte que les québécismes, qui étaient traités sur [un] pied d'égalité [avec] les autres mots du français, n'étaient pas identifiés par une marque particulière. (Poirier 1995 : 12)

Selon Hélène Cajolet-Laganière⁵, qui a mené une consultation auprès d'un large groupe de langagiers québécois pour connaître leurs besoins en matière de dictionnaire général, il s'est avéré que :

plus de 85 % des répondants souhait[ai]ent un dictionnaire général à l'intérieur duquel les mots et les sens employés au Québec seraient distingués des mots et des sens employés dans les autres pays francophones à l'aide d'une marque quelconque. (Cajolet-Laganière 1998 : 67)

Aussi est-elle d'avis que les lexicographes québécois devraient « accorder une attention particulière aux marques topolectales » (*ibid.*).

Même si les expériences lexicographiques du *DFP* et du *DQA* n'ont pas été concluantes, elles ont eu le mérite de poser un certain nombre de questions jusque-là informées, dont celle de l'inclusion des francismes dans un dictionnaire québécois :

Un autre aspect extrêmement novateur du *DFP* et du *DQA* est passé presque inaperçu ou comme allant de soi, au Québec tout au moins : celui de l'inclusion, de la reconnaissance et de l'identification d'un certain nombre de mots et d'emplois considérés comme particuliers au français de France. (Verreault 1996 : 200)

Étant donné le statut international du français de France, il n'est pas étonnant que le public québécois attende d'un dictionnaire général québécois qu'en plus de décrire le lexique français d'usage commun et les principaux québécismes, cet ouvrage répertorie un certain nombre de francismes. Se pose alors la question du choix de ces francismes, de leur traitement et évidemment de leur identification.

1.3 DANS LE DICTIONNAIRE DU PROJET *FRANQUS*

Le tableau 4 fait état du traitement réservé au québécisme⁶ *mitaine* et à son pendant européen, le francisme *moufle*, dans tous les dictionnaires français et québécois dont il a été question jusqu'à présent, ainsi que dans le dictionnaire

4 Par exemple, de la paire synonymique *mitaine* (québécisme) et *moufle* (francisme), seul le premier emploi est marqué sur le plan topolectal (v. l'article *mitaine* ci-dessous). En fait, ce dictionnaire québécois de difficultés reprend globalement l'attitude des dictionnaires généraux faits en France selon laquelle – comme le reconnaît depuis longtemps Alain Rey – « le français de France est assimilé non seulement à une variété locale, mais à la norme unique de la langue » (Rey 1986 : 30). Au fil des rééditions, le *Multi* ne cesse de se développer et, dans l'absence d'un véritable dictionnaire général usuel québécois, en vient à jouer ce rôle auprès d'un large public.

5 H. Cajolet-Laganière codirige le projet *FRANQUS* avec Pierre Martel.

6 Caractéristique du français québécois, cet emploi n'est toutefois pas exclusif à cette variété. Dans le *NPR* (v. le tableau 4), l'emploi est aussi donné comme vieilli en français; dans le *PL*, il est aussi présenté comme helvétisme. Sur son emploi régional en France et sporadique en Suisse, voir le *Dictionnaire suisse romand*.

FRANQUS en cours de réalisation à Sherbrooke. On pourrait s'attarder sur l'ordonnancement des acceptions, sur les procédés définitoires exploités ainsi que sur la mise en relation de ces emplois par le biais des renvois, mais nous nous en tiendrons ici au seul marquage topolectal.

Partant des dictionnaires français, on voit que l'approche reste la même dans le *Multi* : seul le québécisme est marqué. Un changement radical de perspective est en revanche observable dans le *DFP* et le *DQA* : seul *moufle* est précédé d'une marque. Le traitement proposé par le dictionnaire *FRANQUS* introduit une nouvelle perspective où tous les particularismes répertoriés, qu'ils soient caractéristiques de l'usage québécois du français ou caractéristiques de son usage européen, sont identifiés par une marque topolectale, respectivement FQ et FE⁷.




Schéma 3 : Approche descriptive et marquage topolectal dans *FRANQUS*

Comme le montre notre dernier schéma, la nouvelle approche cherche à répondre aux attentes actuelles du public québécois en démarquant ces deux types d'emplois par rapport au fonds commun du français.

Comparant l'article *mitaine* du dictionnaire *FRANQUS* aux articles correspondants du *DFP* et du *DQA*, on remarque également que l'acception « gant laissant à découvert l'extrémité des doigts » n'est marquée comme francisme que dans ces deux derniers ouvrages. Dans le *DFP* et le *DQA*, on a sans doute marqué cet emploi parce qu'il est relativement peu fréquent au Québec, comparativement à l'autre acception. Mais cette basse fréquence ne relève que très indirectement de la variation linguistique : elle s'explique d'abord et avant tout par l'emploi peu répandu au Québec du type de gant en cause. Comme le français québécois ne dispose pas de mot équivalent proprement québécois, on n'est pas vraiment devant un cas de variation topolectale. Dans le dictionnaire *FRANQUS*, le marquage topolectal est réservé aux cas où le français québécois et le français hexagonal peuvent être clairement distingués sur une base strictement linguistique.

⁷ Les emplois français que les Québécois perçoivent comme non caractéristiques de leur variété linguistique correspondent dans une large proportion à des emplois que le français hexagonal partage avec les variétés européennes voisines, d'où l'abréviation FE provisoirement retenue dans le dictionnaire *FRANQUS*.

		mitaine		moufle
<i>NPR</i>	Vx ou région. (Canada)	1. Moufle.		1 Cour. Pièce de l'habillement qui couvre entièrement la main, sans séparation pour les doigts, sauf pour le pouce. => gant , région. mitaine .
		2. Mod. Gant qui laisse à nu les deux dernières phalanges des doigts. [...]		
<i>PL</i>		1. Gant qui ne couvre que les premières phalanges.		
	Québec, Suisse.	2. Moufle.		1. Gant, généralement fourré, où il n'y a de séparation que pour le pouce.
<i>Multi</i>		1. Gant qui découvre les doigts.		
		2. Partie de l'habillement qui couvre la main, sans séparation pour les doigts, sauf pour le pouce. <i>SYNONYME</i> > moufle. <i>NOTE</i> > Ne pas confondre avec le nom gant , partie de l'habillement qui couvre la main et les doigts séparément.		Partie de l'habillement qui couvre la main, sans séparation pour les doigts, sauf pour le pouce. <i>SYNONYME</i> > mitaine. <i>NOTE</i> > Ne pas confondre avec le nom gant , partie de l'habillement qui couvre la main et les doigts séparément.
<i>DFP</i>		1. Gros gant qu'on porte pour se protéger du froid, recouvrant entièrement la main et ne comportant pas de séparations pour les doigts, excepté pour le pouce. [...]	(France)	1. Gros gant ne comportant pas de séparations pour les doigts, excepté pour le pouce. Rem. Usage limité à la langue publicitaire au Québec. V. mitaine.

	(France)	2. Gant de femme qui laisse découvertes les deux dernières phalanges des doigts. [...].		
<i>DQA</i>		1. Pièce d'habillement qui protège contre qqch.		
		1. Gros gant qui recouvre entièrement la main, sans séparation pour les doigts, sauf pour le pouce, porté pour se protéger du froid. => mitasse , 1. moufle . [...]	(France)	Sorte de gant fourré sans séparation pour les doigts sauf pour le pouce. => mitaine .
	(France)	3. Gant, surtout de femme, qui laisse à nu les deux dernières phalanges des doigts. [...]		
<i>FRANQUS</i>		1. Gant laissant à découvert l'extrémité des doigts.		
	FQ	2. Gant chaud sans séparation pour les doigts, excepté pour le pouce. => moufle . [...] REM. Ce sens est vieilli en français européen.	FE	1. Gant chaud sans séparation pour les doigts, excepté pour le pouce. => mitaine . [...]

Tableau 4. Mise en relation de mitaine et de moufle⁸

⁸ Nous ne reproduisons ici que l'essentiel des acceptions relevant de l'habillement de la main. L'emploi caractéristique du français québécois est mis en grisé et toutes les marques topolectales ont été mises en évidence dans une colonne distincte.

1.4 VARIÉTÉS LINGUISTIQUES ET CONTEXTES RÉFÉRENTIELS : LIMITES DU MARQUAGE TOPOLECTAL

Selon qu'il dessert les communautés francophones hexagonale ou québécoise, le français vient s'inscrire dans deux contextes fortement différenciés. Parmi l'ensemble des différences que l'on peut remarquer entre les usages linguistiques de ces deux communautés, deux types généraux doivent être distingués.

La paire *moufle* / *mitaine* vient d'illustrer clairement le premier type, où la différence est essentiellement linguistique, sans lien avec le contexte référentiel. C'est parce qu'on a appris le français en France ou au Québec que l'on a spontanément recours à l'une ou à l'autre de ces dénominations.

Cependant, si l'on va au lycée ou au cégep, ce n'est pas en raison de sa variété maternelle de français, mais en raison du système d'éducation qui a cours dans la communauté où l'on vit. La différence linguistique qui existe entre *lycée* / *cégep* dépend donc directement du contexte référentiel. Même si, sur le plan discursif, ces mots contribuent à caractériser les usages hexagonal et québécois⁹, il ne s'agit pas à strictement parler de francisme et de québécisme. De telles différences, qui relèvent du deuxième type général, ne devraient donc pas, en principe, faire l'objet d'un marquage topolectal. Le recours à un indicateur contextuel (différent de la marque d'usage) nous semble plus approprié.

Ces deux grands types de différences linguistiques sont représentés dans le vocabulaire politique dont il sera dorénavant question dans la suite de cet article.

2. TRAITEMENT DU VOCABULAIRE POLITIQUE : MISE EN RELATION DES USAGES QUÉBÉCOIS ET FRANÇAIS

Comme nombre d'autres sous-ensembles lexicaux étroitement associés à un contexte variable, contexte physique (par exemple, le vocabulaire de la faune et de la flore) ou socioculturel (par exemple, les vocabulaires administratifs et sportifs), le vocabulaire politique québécois présente de nombreuses particularités par rapport à son pendant hexagonal.

	Variation essentiellement linguistique	Variation en lien direct avec le contexte référentiel
Mots caractéristiques	<i>caucus, chefferie, voteur</i>	<i>péquiste, bloquiste, felquiste, gouverneur général, MRC, sous-ministre, tory, whip</i>
Sens caractéristiques		<i>patriote, libéral, arrondissement, gouverneur, canton, préfet</i>

Tableau 5 : *Quelques éléments caractéristiques du vocabulaire politique québécois*

À la base, le régime politique canadien, qui est de type parlementaire et qui comporte deux niveaux de juridiction (fédéral et provincial), se distingue nettement du régime républicain de la France; le Québec, comme province canadien-

⁹ Ce n'était pas vraiment le cas de *mitaine* « gant laissant à découvert l'extrémité des doigts » qui n'a pas son pendant québécois.

ne, possède des institutions, des partis politiques, des fonctions politiques qui lui sont propres, ainsi que des faits historiques qui le caractérisent. Il existe dans une moindre mesure, des différences inhérentes à la langue elle-même.

Le tableau suivant présente quelques éléments caractéristiques du vocabulaire politique québécois en les classant selon leur nature linguistique (mots ou sens) et selon le type de variation en cause (essentiellement linguistique ou en lien direct avec le contexte référentiel).

2.1 CAS DE VARIATION ESSENTIELLEMENT LINGUISTIQUE

À la lecture du tableau, l'on remarque que la colonne des cas de variation essentiellement linguistique est moins garnie que celle consignant les cas de variation en lien direct avec le contexte référentiel. En effet, toutes les différences que nous venons d'évoquer quant à la nature du régime politique canadien, des partis politiques canadiens et québécois, des événements historiques, etc., se reflètent dans le vocabulaire, mais tiennent d'abord du contexte référentiel.

• *Voteur*

Le mot *voteur*, qui est synonyme de *votant*, permet d'aborder un aspect délicat de la description des particularismes québécois, soit la prise en compte de leur évaluation normative. Malgré sa fréquence élevée dans les journaux québécois, *voteur* est absent des dictionnaires consultés pour cette étude, sauf du *Multi* qui le signale comme forme fautive sous *votant*. Cet emploi ne semble effectivement pas relever du français québécois standard. Dans le *Grand Dictionnaire terminologique* (GDT) de l'Office québécois de la langue française, à l'article *votant*, on peut lire la remarque suivante : « Le mot *voteur* n'existe pas. » Or le mot existe bel et bien... Il est attesté 10 fois dans la *Banque de données textuelles de Sherbrooke* (BDTS)¹⁰ et plus de 600 fois dans la banque de données médiatiques *Biblio Branchée* (section canadienne-française, en date du 23 avril 2007)¹¹. Même si la remarque du GDT est contredite par l'usage observé, elle invite à la prudence en ce qui a trait à l'inclusion de ce mot dans un dictionnaire général du français au Québec. On sait que la mauvaise réception du précédent dictionnaire général québécois (*DQA*) est en grande partie due à sa trop grande ouverture aux emplois critiqués ou non standard. S'il est intégré à la nomenclature du dictionnaire en cours de rédaction, *voteur* devra donc être identifié comme tel. La situation est différente pour les deux particularismes linguistiques *chefferie* et *caucus* ; après avoir été critiqués comme des écarts par rapport à la norme décrite dans les dictionnaires du français produits en France, ils sont maintenant passés dans l'usage standard et assez généralement acceptés comme tels¹², comme le démontrent les deux remarques suivantes extraites du GDT :

Employé couramment aujourd'hui (au Québec), le mot « caucus » est un canadienisme dont l'usage se justifie parce qu'il est à la fois utile et fonctionnel ; il

10 Mise sur pied par des chercheurs du CATIFQ, la BDTS réunit aujourd'hui plus de 52 millions d'occurrences provenant de textes de genres divers (littéraires, journalistiques, spécialisés, etc.). (Au sujet de la BDTS, voir le site du projet FRANQUS.)

11 En comparaison, dans le volet européen de *Biblio Branchée*, *voteur* n'apparaît que dans quatre articles.

12 En 1993, le *DQA* faisait encore état de la critique (voir les articles ci-dessous où *chefferie* fait l'objet d'une remarque et où *caucus* est précédé de la marque *Anglic.*, mise pour *anglicisme à proscrire*). Et le *Multi* présente encore l'acception québécoise de *chefferie* comme une impropriété.

dispense de recourir à des périphrases encombrantes et peu adéquates. (fiche 3, 1989)

Le terme *chefferie* fait partie du vocabulaire politique du Canada français depuis plus d'un siècle. Bien qu'il soit encore critiqué dans certains ouvrages correctifs, il est parfaitement conforme au système linguistique du français. (2005)

- *Chefferie*

Le mot *chefferie* est d'emploi courant au Québec dans le sens de « direction d'un parti politique ». Il est attesté dans plus de 1 000 articles de *Biblio Branchée* (section canadienne-française)¹³ et dans 22 documents de la *BDTS*. Voici comment les dictionnaires consultés pour cette étude traitent l'acception du mot qui relève du domaine politique québécois :

<i>NPR</i> 2007		[Absent]
<i>PL</i>	Québec	[...] 3. Direction d'un parti politique.
<i>DFP</i>	—	Direction d'un parti politique. <i>Candidat, course, congrès à la chefferie.</i>
<i>DQA</i>	—	Direction d'un parti politique. <i>Elle annonce sa candidature à la chefferie. Congrès, course à la chefferie.</i> Fam. <i>Se présenter à la chefferie</i> , poser sa candidature. — REM. Cet emploi est critiqué.
<i>FRANQUS</i>	FQ	I. Direction d'un parti politique. <i>Candidat à la chefferie. Course, congrès à la chefferie. [...]</i>

Tableau 6. Traitement du mot *chefferie*

Le *PL* ajoute la marque topolectale *Québec* devant l'acception concernée, celle-ci n'ayant effectivement cours qu'au Québec. Le mot *chefferie* est bien attesté dans le volet européen de *Biblio Branchée* (208 articles), mais dans d'autres emplois faisant généralement référence au chef traditionnel africain¹⁴. Le *DFP* et le *DQA*, conformément à leur politique éditoriale, ne mettent aucune marque topolectale et n'indiquent donc pas à leurs lecteurs qu'il s'agit d'un québécisme. Le dictionnaire *FRANQUS*, quant à lui, apporte cette précision par la marque topolectale FQ. Par ailleurs, on peut noter qu'à l'encontre de l'ordonnancement des acceptions de *chefferie* proposé par le *PL*, les dictionnaires québécois présentent d'abord l'emploi québécois.

- *Caucus*

Caucus est lui aussi de haute fréquence dans le discours québécois (217 attestations dans la *BDTS* et plus de 1 000 documents dans la section canadienne-française de *Biblio Branchée*). Nous commençons avec ce cas à entrevoir comment le contexte référentiel peut influencer la variation géographique.

13 En fait, *Biblio Branchée* cesse de compter après 1 000 textes et s'en tient alors à l'indication : « plus de 1 000 documents ».

14 Nous avons examiné les 100 premiers contextes et rencontré une douzaine d'attestations d'autres emplois marginaux : « direction d'un hôpital » (6 fois) ; « direction (éditoriale) d'un journal » (2 fois) ; enfin, le mot est utilisé pour désigner les dirigeants politiques des États-Unis (2 fois) ou pour parler du contexte québécois (2 fois).

Comme l'indiquent les marques topolectales du *NPR* et du dictionnaire *FRANQUS*, l'emploi généralisé du mot *caucus* au sens de « réunion politique à huis clos » caractérise clairement le français québécois par rapport au français hexagonal où l'emploi de *caucus* est nettement plus limité, comme on va le voir.

	Marque topolectale	Indicateur context. initial	
<i>NPR</i>	Région. (Canada)		Réunion à huis clos (des élus d'un parti politique).
<i>PL</i>		Au Canada et aux États- Unis	Réunion à huis clos des dirigeants d'un parti politique ; personnes ainsi réunies.
<i>DFQ</i>			1. Réunion à huis clos où les parlementaires d'un même parti politique discutent de la conduite, de la stratégie, etc., de leur parti. <i>Députés réunis en caucus. Président du caucus. Par méton.</i> Ceux qui participent à cette réunion.
<i>DQA</i>			Anglic. 1. Réunion à huis clos au cours de laquelle des députés ou la direction d'un parti politique discutent des affaires du parti. <i>Le caucus des péquistes. La présidente du caucus.</i> Ensemble des personnes ainsi réunies. <i>Le caucus se prononcera demain sur l'avortement.</i>
<i>FRANQUS</i>	FQ		1. Réunion à huis clos des élus ou des parlementaires d'un parti politique ; personnes ainsi réunies. <i>Caucus hebdomadaire, spécial. Caucus régional.</i> Le président du caucus. <i>Convoquer une réunion du caucus.</i> Un débat au sein du caucus. « <i>Ces nouveaux députés [...] arboraient comme un étendard un air de confiance et d'émulation qui les fit remarquer dès leur premier pas dans les caucus et à la Chambre</i> » (G.-É. Lapalme, 1969). REM. : Aux États-Unis, <i>caucus</i> désigne également une réunion de campagne électorale où les membres d'un parti politique choisissent un candidat ou des délégués.

Tableau 7. Traitement du mot *caucus*

Du côté du *PL*, on constate qu'au lieu de la marque topolectale et contrairement à ce qu'il avait fait sous *chefferie*, ce dictionnaire opte ici pour un indicateur contextuel intégré en début de définition. Est-ce que cela revient au même? Pas tout à fait. En principe, le recours à la marque implique que la restriction d'emploi porte sur le mot lui-même, alors que le recours à l'indicateur contextuel indique que le référent décrit est associé à un contexte géographique restreint. Le traitement proposé par le *PL* est ambigu : comment interpréter cet indicateur si le reste de la définition ne décrit pas un référent spécifique à l'Amérique du Nord? Ce cas est intéressant puisqu'il permet de faire le pont entre les cas de variation essentiellement linguistique qui précèdent et les cas suivants où la variation est en lien direct avec le contexte référentiel.

Si le *PL* réfère aux États-Unis, c'est que dans le discours hexagonal, *caucus* ne s'emploie de façon générale qu'en référence au contexte politique des États-Unis et dans un sens qui ne correspond pas exactement à l'emploi québécois. Alors qu'au Canada, le caucus est la réunion à huis clos des députés d'un parti politique (des députés et des sénateurs dans le cas du caucus d'un parti fédéral), du côté de nos voisins du Sud, le mot *caucus* s'emploie aussi pour désigner de vastes réunions politiques où les membres d'un parti se prononcent sur la sélection de délégués ou de candidats¹⁵. Or, dans le volet européen de *Biblio Branchée*, près de 70 % des 676 documents attestant le mot *caucus* font référence au contexte politique étatsunien¹⁶, par exemple dans « *le caucus de l'Iowa* » ou dans *(Congressional) Black caucus* (nom du collectif des élus noirs du Congrès américain). Aucune occurrence du mot ne semble faire référence aux contextes politiques canadien ou québécois¹⁷. L'article du *PL* gagnerait de toute évidence à être revu pour mieux rendre compte de la variation des usages entre le français québécois et le français hexagonal, ou si l'emploi québécois n'est pas retenu, pour mieux rendre compte du contexte où les médias européens optent pour *caucus* plutôt que pour *huis clos*.

2.2 CAS DE VARIATION EN LIEN DIRECT AVEC LE CONTEXTE RÉFÉRENTIEL

Les cas de variation les plus fréquents dans le corpus politique sont en lien direct avec un changement de contexte référentiel. Une différence de contexte entre le Québec et la France ne peut manquer de susciter l'apparition dans les discours québécois et français de mots et de sens différents. C'est ce qui explique, par exemple, l'emploi courant de *péquiste* (du Parti québécois) dans les médias québécois et de *frontiste* (du parti du Front national) dans les médias français.

15 C'est cet emploi que l'*American Heritage Dictionary* décrit dans son acception 1a ("A meeting of the local members of a political party especially to select delegates to a convention or register preferences for candidates running for office."); son acception 1b correspond davantage à l'usage québécois ("A closed meeting of party members within a legislative body to decide on questions of policy or leadership").

16 D'après une recherche croisée avec *caucus* et *États-Unis*.

17 D'après quelques recherches croisées, notamment avec *caucus* et (*Québec* ou *Canada*). Par ailleurs, le journal belge *Le Soir* présente quelques occurrences du sens de « huis-clos politique » en lien avec le contexte européen : « se sont alors succédé de multiples *caucus* dans les couloirs du parlement [européen] » (jeudi 1^{er} décembre 2005, p. 4).

	Marque topolectale	Indicateur contextuel initial	Possibilité d'autres indicateurs intégrés à la définition
<i>NPR</i>	Région. (Canada)		Haut fonctionnaire responsable de l'administration d'un ministère. [...]
<i>PL</i>		Au Canada,	haut fonctionnaire qui seconde un ministre.
<i>FRANQUS</i>		Au Canada,	haut fonctionnaire auquel un mi- nistre confie la direction de son ministère.

Tableau 8. Traitement du mot sous-ministre

Comme il ne s'agit pas à proprement parler de particularismes linguistiques, les articles de dictionnaires portant sur ces mots ne devraient pas en principe comporter de marque topolectale ; c'est le contexte référentiel qui doit être précisé (en fonction des besoins du public visé). Les exemples suivants de *sous-ministre* et de *gouverneur général* montrent que, dans les dictionnaires français actuels, la distinction entre variation strictement linguistique (signalée par la marque topolectale) et variation d'abord référentielle (signalée par l'indicateur contextuel) ne reçoit pas de traitement uniforme (comme nous avons commencé à le voir avec le mot *caucus*).

	Marque topolectale	Indicateur contextuel initial	Possibilité d'autres indicateurs intégrés à la définition
<i>NPR</i>		Au Canada,	représentant(e) de la reine (ou du roi), nommé(e) pour cinq ans par le souverain d'Angleterre, sur la recommandation du Premier mi- nistre du Canada. [...]
<i>PL</i>		Au Canada,	représentant de la reine (ou du roi) d'Angleterre.
	Belgique		Fonctionnaire inamovible placé à la tête d'une province.
<i>FRANQUS</i>			Chef d'État et commandant en chef du Canada, représentant de la Couronne britannique, nommé pour cinq ans. [...]

Tableau 9. Traitement du mot gouverneur général

- *Sous-ministre et gouverneur général*

Sous-ministre et gouverneur général sont deux titres de fonction liés au contexte politique canadien. Comme on le voit (Tableaux 8 et 9), le *NPR* accompagne le premier titre d'une marque d'usage et définit le second en commençant par un indicateur contextuel.

De son côté, le *PL* a recours à l'indicateur contextuel dans les deux cas où il est question du Canada, mais à l'article *gouverneur général*, on voit qu'il met

la marque *Belgique* devant l'acception suivante décrivant un titre de fonction relevant du contexte belge...

CONCLUSION

Après L.-A. Bélisle, après les rédacteurs du *DFP* et du *DQA*, ce sont maintenant les rédacteurs du projet *FRANQUS* qui travaillent à l'élaboration d'un dictionnaire général du français à l'intention des Québécois. On a vu que, depuis le milieu du *xx*^e siècle, l'approche descriptive mise en avant par les lexicographes québécois a connu des changements importants, qui ont eu un effet direct sur le marquage topolectal des emplois caractéristiques de l'usage québécois ou de l'usage hexagonal. À l'heure où les dictionnaires usuels du français ouvrent de plus en plus leurs pages aux mots et emplois caractéristiques des diverses communautés francophones, il nous semble important de mieux départager et identifier les deux grands types de variation que nous venons d'illustrer à partir du vocabulaire politique surtout, mais qui cohabitent dans la plupart des grands champs thématiques. Les outils lexicographiques appropriés sont déjà disponibles : marque d'usage pour les cas de variation essentiellement linguistique et indicateur contextuel pour les cas où la différence d'usage observée dépend directement d'une différence de contexte référentiel. Il ne reste qu'à les exploiter de façon plus systématique; c'est ce que les rédacteurs du projet *FRANQUS* se proposent de faire.

RÉFÉRENCES BIBLIOGRAPHIQUES

- American Heritage Dictionary of the English Language (The)*, 2000, fourth edition; <http://www.bartleby.com/61/5/C0170500.html> (consulté le 04.05.07).
- BEAUJEAN Amédée, 1875, *Dictionnaire de la langue française. Abrégé du dictionnaire de É. Littré*, Paris, Hachette, iij - 1 294 p.
- BÉLISLE Louis-Alexandre, 1957, *Dictionnaire général de la langue française au Canada*, Québec, Bélisle éditeur, [xiv] - 1 390 p. ; 1971, 2^e éd. [xvj] - 1390 p. ; 1974, édition spéciale, Québec et Montréal, Bélisle et Sondec, [xvj] - 1 487 p. ; 1979, 3^e éd., *Dictionnaire nord-américain de la langue française*, Montréal, Beauchemin, [xiv] - 1 196 p.
- CAJOLET-LAGANIÈRE Hélène, 1998, « Attentes et besoins du public québécois en matière de dictionnaires de la langue », dans Louis Mercier et Claude Verreault (éds), *Les Marques lexicographiques en contexte québécois*, Québec, Gouvernement du Québec, Office de la langue française, p. 61-70.
- DE VILLERS Marie-Éva, 2003 (4^e éd.), *Multidictionnaire de la langue française*, Montréal, Éditions Québec Amérique, xxv - 1 542 p. ; 1988, 1^{re} éd., *Multidictionnaire des difficultés de la langue française*. [Multi]
- Dictionnaire du français plus : à l'usage des francophones d'Amérique*, 1988, Montréal, Centre éducatif et culturel, 1 856 p.
- Dictionnaire québécois d'aujourd'hui*, 1993, 1^{re} édition 1992, sous la direction de Jean-Claude Boulanger et la supervision d'Alain Rey, Saint-Laurent, Dico-robert, xxxviii - 1 274 p. (partie langue).
- Dictionnaire suisse romand. Particularités lexicales du français contemporain*, 1997, conçu et rédigé par André Thibault sous la direction de Pierre Knecht, Genève, Éditions Zoé, 854 p.

FRANQUS (site du projet) : <http://franqus.usherbrooke.ca>.

Grand Dictionnaire terminologique (en ligne), élaboré et produit par l'Office québécois de la langue française [et] Semantix
<http://www.granddictionnaire.com> (consulté le 04.05.07).

MERCIER Louis, 2002, « Le français, une langue qui varie selon les contextes », dans Claude Verreault, Louis Mercier et Thomas Lavoie (éds), *Le Français, une langue à apprivoiser : textes des conférences prononcées au Musée de la civilisation (Québec, 2000-2001) dans le cadre de l'exposition Une grande langue : le français dans tous ses états*, Québec, Les Presses de l'Université Laval, p. 41-60.

Nouveau Petit Robert, 2006 (sous la direction de Josette Rey-Debove et Alain Rey), Paris, Dictionnaires Le Robert, xlij – 2 837 p.

POIRIER Claude, 1995, « De la soumission à la prise de parole : le cheminement de la lexicographie au Québec », dans B.B. Kachru and H. Kahane (eds.), *Cultures, Ideologies and the Dictionary, Studies in Honor of Ladislav Zgusta*, Tübingen, Niemeyer, p. 237-252.

REY Alain, 1986, « La variation linguistique dans l'espace et les dictionnaires », dans Lionel Boisvert, Claude Poirier et Claude Verreault (éds), *La lexicographie québécoise, bilan et Perspectives, Actes du colloque organisé par le TLFQ à l'Université Laval les 11 et 12 avril 1985*, Québec, Les Presses de l'Université Laval (« La Langue française au Québec », 8), p. 23-36.

VERREAULT Claude, 1996, « Inclusion, reconnaissance et identification des francismes dans les dictionnaires québécois : problèmes et méthodes à la lumière de l'expérience du *Dictionnaire québécois d'aujourd'hui* », dans *Français du Canada, français de France : actes du 4^e colloque international de Chicoutimi, Québec, du 21 au 24 septembre 1994*, Tübingen, Niemeyer (« Canadiana Romanica », 12), p. 199-208.

CONTOURNER LA MARQUE DIATOPIQUE : VITALITÉ DES EMPLOIS ET PRÉSENCE DES TOPOLECTISMES DANS LES DICTIONNAIRES DE VARIÉTÉS DE FRANÇAIS

par Jean Nicolas DE SURMONT

« À bien prendre les choses, le dictionnaire est le livre
par excellence. Tous les autres livres sont dedans : il ne
s'agit plus que de les en tirer. »

Anatole France

Résumé — La lexicographie du français de Belgique a suscité depuis peu un certain intérêt, peut-être davantage de la part des lexicographes que de la part des métalexicographes. Notons par exemple les récents dictionnaires de Christian Delcourt, de Georges Lebouc puis de l'inclusion plus importante de belgicismes dans le *Petit Larousse illustré* grâce à la collaboration Jean-Marie Klinckenberg. Le présent article s'intéressera au marquage diatopique, plus particulièrement celui des belgicismes, en ciblant davantage l'observation sur les divergences de traitement et de sélection dans quelques dictionnaires récents de langue française (notamment le *Dictionnaire Hachette*, le *Petit Larousse illustré*, le *Grand Robert électronique*, la BDLP [*Base de données lexicographiques pan-francophones*]). Différents problèmes surgissant lors de la sélection de la nomenclature de lexies régionales seront posés : par exemple l'omission de certains lexèmes du français standard dans les dictionnaires du français standard les assimilant ainsi à une pratique régionale, la vitalité relative, en chute ou non représentative de certains belgicismes, le fait que l'on découvre un emploi dans une autre variété nationale de français (cela ne lui retire pas pour autant son statut de toplectisme mais en revanche il cesse dès lors d'être un belgicisme). Ainsi la vitalité des francophonismes n'est pas forcément à poser comme critère de représentativité d'une nomenclature périphérique dans un dictionnaire standard mais cela doit tenir aussi compte du fait que les pratiques d'inclusions et de marquages varient d'un dictionnaire à un autre. À partir d'un corpus de belgicismes nous étudierons les problèmes que soulève l'étude des belgicismes, de leur marquage, de leur vitalité, de leur sélection et de leur traitement lexicographique.

Mots clés — belgicisme, marque topolectale, marque diastratique, discours lexicographique, enquête de vitalité.

Abstract — Labels that refer to places, and in particular labels that refer to Belgian French, are the focus of this article. First we point out the different types of lexicographical treatments and then examine the various treatments of Belgian French in several recent French dictionaries such as the *Hachette Dictionary*, the *Petit Larousse illustré*, the *Grand Robert électronique*, and the *BDLP*. Different problems emerged during this study, including: which lexical items are part of the entry and which are not; what are the criteria used to select the Belgian French items to be included in a standard French dictionary. We argue that from among the selection criteria, the vitality of a word is important, as is to make sure that so-called Belgian usage is really confined to Belgium and that the word is not used in other parts of the French speaking world.

Keywords — Belgian French, labels, sociolects, lexicographical discourse, word's vitality.

BILAN DE LA QUESTION

Le marquage a fait l'objet de travaux récurrents depuis le début des années 1980, notamment ceux d'Henri Béjoint (1981a et 1981b) sur les dictionnaires de langue anglaise, de Chantal Girardin (1987) sur les aspects sociolinguistiques des dictionnaires culturels français, de Pierre Corbin (1989) sur les marques stylistiques et diastratiques, de P. A. Messelaar (1988) sur les marques *familier* et *populaire* et surtout de Michel Glatigny à qui l'on doit une coordination d'ouvrage (1990) et une monographie (1998) sur les marques d'usage contemporain et enfin d'un ouvrage collectif dirigé par Louis Mercier et Claude Vervault (1998) sur les marques d'usage en contexte québécois. L'ensemble des articles et monographies en langue française sur les dictionnaires de langue française montre qu'un intérêt plus important semble avoir été accordé aux marques diastratiques et diaphasiques plutôt qu'aux marques diatopiques, outre le fait que la lexicographie belge reste presque inexplorée. Dans la poursuite de notre contribution à l'étude des marques diatopiques présentée au Congrès Euralex à Lorient en 2004, nous allons proposer une réflexion plus générale sur le rapport entre la marque et la technique lexicographique en soulevant différents problèmes liés à l'inclusion de topolectismes dans les dictionnaires.

CONSIDÉRATION GÉNÉRALE SUR LA MARQUE D'USAGE

Avant de nous intéresser à proprement parler à la marque diatopique, il convient de rappeler certains faits généraux sur la marque d'usage. La citation d'Anatole France en exergue provient de l'entrée *dictionnaire* du *Grand Robert de la langue française* (dorénavant *GR*). Elle caractérise bien la dialectique entre l'étude de terrain et le dépouillement du corpus linguistique (« les autres livres ») et l'élaboration de la nomenclature dictionnaire d'une part et l'activité d'encodage d'autre part. Mais elle nous laisse sur notre faim quant aux filiations entre dictionnaires de même tradition éditoriale ou de tradition éditoriale différente (voir De Surmont 2005), entre dictionnaires généraux monolingues (doré-

navant DGM) et dictionnaires de variétés de français (dorénavant DVF)¹ cela pourrait bien nous laisser croire que les dictionnaires de toponymies ne sont pas tous dans les DGM pour paraphraser Anatole France. Ce fait pourrait être intéressant pour signaler les liens potentiels dans l'établissement de la nomenclature et du marquage entre dictionnaires d'obédience variée.

Avec la typographie, les signes de ponctuation, l'appareil diacritique et le protocole de numérotation des sens, les marques d'usage sont l'un des cinq types de ressources du lexicographe dans l'élaboration de la microstructure (Boulanger, 2001 : 248). Le fait de consigner une unité lexicale dans un dictionnaire procède d'un choix dictionnaire, d'une sélection judicieuse et pas toujours facile pour un lexicographe. Le simple fait d'accepter un mot au sein de la nomenclature pourrait nous laisser croire que le mot-vedette est fréquent. Sur cette base, il est généralement admis que l'article lexicographique comme tel permet d'évaluer la fréquence d'un lexème, la fréquence de son sens par sa structuration ou ses rubriques. Par le recours à des marques d'usage, le rédacteur classe l'unité à l'intérieur d'un champ disciplinaire ; par une marque de domaine, dans le champ discursif par la marque diastatique, etc. ; par le recours à la citation, le lexicographe pose comme modèle des emplois à la manière d'une anthologie littéraire. Les rubriques « Citations » et les marques diastatiques sont respectivement liées au discours sur la littérature et au discours sur la langue, alors que les marques de domaine sont fortement en relation avec l'avancement des sciences à une époque donnée². Les premières relèvent de l'activité d'un choix dictionnaire qui évalue la littérarité des auteurs et des contextes en même temps que leur adéquation avec le sens consigné, alors que les marques diastatiques et de domaines relèvent de la qualification de la représentation collective des registres de langue par le lexicographe et de l'épistémologie contemporaine.

Les marques diatopiques portent un jugement sinon une observation sur la notion de régional, de centre et de périphérie. La description du français de référence est ainsi stratifiée par différentes marques qui rendent compte de l'hétérogénéité des pratiques linguistiques au sein de l'unité de la langue. Toutes ces formes d'appréciations ne sont pas éternelles mais sont inscrites dans un contexte social, une *pragmatique lexicoculturelle* pour reprendre le terme forgé par Robert Galisson (1999) puis porté à l'honneur par Jean Pruvost (2005 : 7-10). Elles diffèrent également selon le point de référence choisi dans un dictionnaire généralement en partage dans les différents pays de la Francophonie.

MARQUAGE DES BELGICISMES

Dans un précédent travail sur la marque lexicographique je m'étais intéressé aux différents types de marques diatopiques en prenant notamment l'exemple du

1 En effet inclure des particularités du français extra-hexagonal dans un DGM consiste en quelque sorte à « normaliser la régionalité », comme l'affirme adéquatement Didier de Robillard. En revanche, ériger un français périphérique comme une norme en marquant l'usage du français de référence par « francisme » équivaut à « régionaliser la norme » (1993 : 152).

2 Pour cette raison nous croyons que la prise en compte d'enquêtes de vitalité permet de rendre compte de l'évolution et du maintien des emplois. Il faut de toute manière admettre que la valeur sociolinguistique des variétés de français non-hexagonales est toujours plus ou moins en décalage avec l'usage du fait de l'absence fréquente de réédition des dictionnaires de toponymies. Notons aussi que les pratiques mêmes changent avec le temps comme nous le signalait lui-même Claude Poirier en interview (Source : Interview avec Claude Poirier, Radio Basse Ville Québec, 10 février 1993).

marquage des belgicisms. À ce niveau, l'activité de rénovation du lexique varie d'un dictionnaire à l'autre. Ainsi, entre la parution du premier recueil de cacographies belge en 1806, signé de la plume d'un français, A.-F. Poyart, et la prise en compte des belgicisms par le *Petit Robert* (dorénavant PR) à la fin des années 1970 grâce à la collaboration du spécialiste de littérature wallonne Maurice Piron et d'Emile Seutin, il y avait déjà eu en Belgique une trentaine de recueils de cacographies, dont une petite douzaine au XIX^e siècle. Nous avons notamment signalé dans cette contribution l'absence d'ajouts de belgicisms par le PR depuis le décès de Maurice Piron³ et établi notamment les lacunes de la nomenclature des belgicisms, par la confrontation entre le corpus de Valibel et les belgicisms consignés dans le *Grand Robert* (dorénavant GR).

Il semble clair que dans les années 1970 les éditions Robert avaient profité de l'essor du mouvement d'inclusion des topolectismes dans les DGM pour faire de même, ce qui ne s'est pas avéré une pratique régulière au fil des ans, contrairement au *Petit Larousse Illustré* (dorénavant *PLI*)⁴.

En effet, le *PLI* présente des mots, des sens et des emplois nouveaux, des locutions nouvelles et de nouveaux noms propres ; par exemple : 305 en 1977, 100 en 1987, environ 3 000 pour la refonte importante de 1989⁵, 174 en 1996, 148 [dont 37 noms propres] en 1997. Parmi ceux-ci il faut signaler que l'ajout de belgicisms se chiffre entre 2 et 10 à chaque millésime. Toute proportion gardée l'ajout, de belgicisms est annuellement inférieur à celui de québécoismes et est moins représenté dans l'ensemble du *PLI* d'après la révision complète des trois premières lettres de l'alphabet que nous avons faite en 2004 et qui donnait ceci : 148 helvétismes, 100 québécoismes, 86 belgicisms, 20 africanismes, 10 entrées pour les Antilles, 6 réunionismes, 5 pour la région du Nord, 4 pour la Savoie et la Provence, 4 entrées pour la Louisiane, 4 algérianismes, 3 acadianismes, 2 entrées du français régional de l'Ouest, 1 pour la Nouvelle-Calédonie, 1 pour le Nord-Est, 1 pour la Polynésie, 1 pour le Centre, 1 pour l'Alsace⁶. Un simple bilan des ajouts lexicaux et sémantiques belges dans de récentes éditions du *PLI* donne les résultats indiqués plus bas⁷. Notons que pour

3 Nous avons notamment pris contact avec Danièle Morvan afin de lui proposer la collaboration du groupe de recherche Valibel au sein duquel nous travaillions alors. Dans le cadre de ce projet nous avons effectivement analysé en détail la présence des belgicisms dans le *Grand Robert de la langue française*, et dans d'autres sources dictionnaires, travail qui a donné lieu à la révision entière des belgicisms dans le millésime 2007 du *Petit Robert*.

4 On observera en outre que si l'on assimilait les occurrences de « belgicisme » dans le *Grand Robert électronique* comme seule manière de trouver un item belge dans le dictionnaire nous arriverions à 18 occurrences, ce qui est en revanche peu représentatif de l'ensemble des belgicisms du GR.

5 Mentionnons que c'est dans cette refonte que l'éditeur Larousse évoque pour la première fois dans une préface du *Petit Larousse* « la francophonie hors de France ». Voir à ce sujet Yves Garnier, 2005 : 225.

6 Le *Dictionnaire universel Hachette* consigne environ 280 belgicisms lexicaux auxquels il faut en ajouter soixante autres pour les entrées à double, triple, quadruple ou quintuple acceptions, que nous devons à Michel Francard avec la collaboration de Geneviève Geron et quelques autres lexicographes (Source VALIBEL, [s.d], [document tapuscrit], 32 p.) Ils représentent assez fidèlement les particularités belges bien qu'on y trouve plusieurs *topolectismes de partage*, ce qui peut revenir à se poser la question suivante : une particularité régionale reste-t-elle une particularité lorsqu'elle est partagée par une, deux, voire trois communautés linguistiques non hexagonales ? Mentionnons néanmoins que *W.C.*, consigné comme belgicisme dans le travail préparatoire du *Dictionnaire Hachette* par Michel Francard, ne nous semble pas plus belge que le caractère administratif voire policier proprement belge d'*acter*. Il pourrait en fait s'agir selon Francard d'un terme de droit administratif, alors que son champ d'application en droit est plus large. Yves Garnier fait la liste des marques diatopiques du *PLI* dans 2005 : 222-223.

donne les résultats indiqués plus bas⁷. Notons que pour les années 1998 et 2002 nous ne mentionnons que les propositions d'ajouts de Jean-Marie Klinkenberg, omettant les propositions de suppressions qui, en 2002, sont tout aussi importantes. Les propositions du romaniste mettent parfois plusieurs années avant d'être consignées par Larousse :

- 1987 Rien de neuf à signaler du côté des belgicisms (sur 100 ajouts).
- 1996 Seul l'ajout de *covoiturage* est à signaler (sans marque diatopique).
- 1997 Introduction de deux belgicisms : *rattachiste* et *pensionné*.
- 1998 Proposition de Klinkenberg : *accises, baes, bloque, brol, caillant, casserole, dix-heures, élocution, horodateur, krieg, noir, peter*.
- 2002 Proposition de Klinkenberg : *à la rue, agrégation, agrégé, après (par), avaloir, babelutte, barrière nadar, bazar, beau (s. v.), belle, blanche, bleu, bloque, brûlant, cache poussière, cagoulé, carnet, chercher, chou, cloche, communautaire, communautariser, coûter, croustillon, diesel, école, entité, être glosé, escavèche, humanité, pl. humanités, indexer, intercommunale, lavoir, moins, nic-nac ou nicnac, pils, plaine de jeux, priorité, quart, région, régionaliser, régionalisation, résonner, sonner, sous-verre*.
- 2003 Introduction de *bel-étage, logopède, pécule, séniorie (seigneurie) et toutes-boîtes*. Face à cette stagnation le *Dictionnaire universel Hachette* est venu combler un manque en publiant près de 489 belgicisms⁸.
- 2004 Ajout de *communautaire, communautariser*⁹, *coulon, école, humanités, plaine, région, resonner, sonner*.
- 2005 Ajout de *après*¹⁰, *avaloir, bazar, bleu, bloque, brûlant, cloche, intercommunale, nicnac, nic-nac, quart, unifamilial*.
- 2007 Introduction de *accoucheuse*¹¹, *bidon*¹², *boîte, brique, cuisine-cave*.

Si l'on peut expliquer la faible représentation des belgicisms par rapport aux québécoisismes et aux helvétismes dans les DGM par le fait que les Belges ne se sont pas dotés d'une description lexicographique très documentée de leur lexique dans les dernières années ou par le nombre de belgicisms moins élevé, on peut néanmoins reconnaître que la majorité des DGM établissaient leur nomenclature en ne s'appuyant pas sur une enquête de vitalité pan-nationale, alors que

7 Ici encore, le français du Québec est mieux documenté puisque la question de l'inclusion des canadianismes et des québécoisismes dans le *PLI* a déjà été étudiée par Jean-Yves Dugas et Jean-Claude Boulanger notamment.

8 L'écart est encore important entre les belgicisms et les québécoisismes et acadianismes. En effet Claude Poirier informe qu'il a rédigé 1 700 textes lexicographiques pour Hachette, mentionnant que la lexicographie francophone fait à ce moment « un bond en avant » (Claude Poirier, lettre personnelle du 20 août 1998).

9 *Communautariser les soins de santé* dans le journal *Le Soir* du jeudi 26 février 1998, *communautariser le droit d'asile* (1998), *il ne fallait pas communautariser* (2002).

10 « *Par après*, au début du mois d'août, ils pourront entamer le pavage de l'avenue de la Toison d'Or ». (M.-L. Gabriel, *Le Soir*, 14 juillet 1998, p. 18).

11 Bien qu'*accoucheuse* soit défini comme « sage-femme », il apparaît que l'*accoucheuse*, surtout attestée à Bruxelles, soit une infirmière-accoucheuse et non une sage-femme ne pratiquant pas en milieu hospitalier.

12 La majorité du temps attesté avec le verbe *arranger* en cotexte : *Arrange les bidons !*

le romaniste Jean-Marie Klinkenberg a « pu bénéficier, pour la préparation de la refonte 1998, des résultats (inédits) d'enquêtes permettant de se faire une idée exacte de la diffusion des belgicisms, en termes géographiques et en termes de couches d'âge. Ceci a abouti à un véritable nettoyage. »¹³ L'accès à cette enquête de vitalité a donc encouragé la suppression de certains belgicisms comme en fait mention Klinkenberg et l'ajout de plusieurs belgicisms en décembre 2002 dont plusieurs n'ont été publiés que dans le millésime de 2004 ou 2005. Il écrit à ce sujet :

On trouvera toutefois ci-après un nouvel ensemble de propositions de suppression. Il s'agit de termes

tombant en désuétude,

utilisés seulement par les couches les plus âgées de la population,

– d'un usage trop sectoriel,

– utilisés dans une aire géographique restreinte,

dont le référent a disparu. (2002 : f. 1)

On peut néanmoins constater deux lacunes dans l'ajout des belgicisms : les contraintes d'espace empêchent souvent la présence d'unités polylexématiques ou de collocations comme c'est le cas pour la description du français de référence. Le résultat des dépouillements de corpus linguistiques en Belgique, que ce soit pour le dictionnaire de Georges Lebouc (2006) ou le *Dictionnaire du français de Belgique* – en cours d'élaboration au Centre de recherche Valibel – auxquels nous avons collaboré avec Geneviève Geron, Régine Wilmet et Michel Francard, montre des faiblesses au niveau de l'exemplification, ce qui prouve l'importance du travail qui reste à faire en lexicographie variationnelle belge quelques années après le début de l'élaboration de corpus oraux et écrits dans les années 1990. Il reste donc à faire en Belgique la présentation des résultats d'une enquête de vitalité qui validerait sociolinguistiquement l'établissement de la nomenclature. Cette manière de faire différerait de celle du *Dictionnaire des régionalismes de France* de Pierre Rézeau où l'enquête de vitalité est postérieure à l'établissement de la nomenclature. En Belgique, cette enquête de vitalité, qui se fait bien entendu à l'échelle d'une région et non d'une localité¹⁴, permet ainsi l'affinement des marques diachroniques et diatopiques, car les emplois ne sont pas toujours attestés dans l'ensemble de la Wallonie ou de la Belgique. À cet égard, la faible prise en compte des dictionnaires de dialectes bruxellois (à titre d'exemple le Louis Quievreux : 1973), pourrait laisser à croire que la dialectologie wallonne est plus vivace.

Encore faut-il souligner que lors d'une enquête effectuée par Ipsos pour le compte de la maison Larousse en août 2004, les résultats ont montré que 44 % des dictionnaires détenus par des particuliers ont plus de dix ans. Cela signifie donc que les ajouts à la nomenclature – même s'ils prétendent être le plus près possible de la réalité de la langue qu'ils décrivent – ne sont réellement connus du public que bien plus tard¹⁵, sans compter que certains belgicisms n'accèdent

13 Source : Document tapuscrit de Jean-Marie Klinkenberg, Note générale pour la proposition [au PLI] des nouveaux belgicisms soumise en décembre 2002, f. 1.

14 Il s'agirait alors de *subrégionalisme* pour reprendre le métaterme de Pierre Rézeau dans le *Dictionnaire des régionalismes* de France.

15 Sondage Ipsos culture, 5 et 6 août 2004.

pas forcément à la reconnaissance au moment de l'enquête de vitalité, mais sont attestés dans les journaux et seront peut-être fort répandus une fois le travail publié. Pour cette raison nous croyons que la prise en compte d'enquêtes de vitalité permet effectivement de rendre compte de l'évolution et du maintien des emplois. Il faut de toute manière admettre que la valeur sociolinguistique des variétés de français non hexagonales est toujours plus ou moins en décalage avec l'usage, du fait de l'absence fréquente de réédition des dictionnaires de topolécismes¹⁶.

LES OMISSIONS ET L'ENQUÊTE DE VITALITÉ

À partir du moment où un lexème figure dans un DVF, il est susceptible d'être repris dans un DGM. Dans le premier cas la marque ne se justifie pas, mais en revanche l'inclusion est souvent basée sur une pratique déclarée par l'informateur et pas forcément constatée par l'enquêteur¹⁷. Donc l'enquête sur la vitalité d'une lexie se trouve en partie biaisée par cette variable et ainsi en va-t-il de la dispersion ou la diffusion du belgicisme. On peut dès lors critiquer l'adéquation entre la pratique déclarée et la nomenclature.

À l'inverse peut se poser le problème de consigner dans un dictionnaire de topolécismes belges une unité lexicale écartée des dictionnaires français parce qu'ayant échappé à la vigilance des dictionnistes, mais connue en France et en Belgique¹⁸. C'est le cas de la locution *encore un peu*, courante en Belgique probablement grâce à l'influence dialectale mais ne comportant pas un caractère régional. Consignée par Joseph Hanse dans les éditions de 1987 et 1994 de son *Nouveau dictionnaire des difficultés du français moderne* et toujours dans la quatrième édition co-signée par Daniel Blampain comme emploi à « éviter » sous l'entrée *encore* (2000 : 228) et sous *peu* : « Éviter de dire, comme en wallon : [*Encore un peu je tombais*] pour *Un peu plus, je tombais* ou *J'ai failli tomber* » (2000 : 441). La critique de Hanse s'inscrit résolument dans une pensée guidée par l'appartenance des lettres de Belgique à la littérature française¹⁹.

L'emploi est pourtant connu en France, dans des emplois rares mais attestés au milieu du XIX^e siècle chez Honoré de Balzac par exemple :

Encore un peu, je croirais que vous voulez me voler votre dernière maladie et mourir d'une autre main que de la mienne. (1842 : 78)

ou aussi :

Encore un peu, nous allons être de vieux amis.

— Quoique l'amitié doive être près de vous un sentiment peu vulgaire, dit Rastignac, je ne veux jamais être votre ami. (1843 : 156)

16 Les années 1960 marquées par le développement de l'informatique, y compris dans la lexicographie, ont été conscientes des aspects évolutifs rapides du lexique dus aux changements sociaux. L'ouvrage de Dubois et Dubois en a particulièrement rendu compte (1971 : 127 et suiv.)

17 C'est l'interférence de jugements de normalité et d'hypercorrection qui biaisent les pratiques réelles.

18 On peut aussi mentionner le fait que certains belgicismes sont traditionnellement écartés par la lexicographie belge elle-même comme *simoniser* au sens de lustrer, encaustiquer, *fromage d'abbaye*, *mousse-line*, etc.

19 En effet, comme le souligne Jean-Marie Klinkenberg, ce qui doit être mis en relief dans une vision centripète du corpus belge, c'est plutôt l'importance de la contribution belge au patrimoine littéraire français. Elle est implicite dans l'*Histoire illustrée des lettres françaises de Belgique* de Gustave Charlier et Joseph Hanse, Bruxelles, Renaissance du livre, 1958.

L'emploi international d'un toplectisme au sein de plusieurs communautés linguistiques périphériques relève de ce que j'ai nommé un *toplectisme de partage*. Il en va ainsi de *déjeuner* attesté dans le sens belge non seulement au Québec mais également en Suisse et dans le Nord-Est de la France, de *paradis* comme partie supérieure du cadre dans un jeu de *marelle* également omis des dictionnaires français²⁰. Prenons un autre exemple issu du *PR*, dans son édition 2007. Le *PR* innove en consignait un double marquage pour le *toplectisme de partage* « sous plat » attesté au Québec comme en Belgique sans donner d'exemples. On s'étonnera que l'extension géographique soit différente dans les récentes éditions du *PLI* qui ne fait que le mentionner comme belgicisme.

Il en va de même de certains mots français standard récents comme *région* (sens 2), *régionalisation* et *régionaliser* qui, bien qu'introduits récemment comme belgicismes dans le *PLI*, sont d'emplois assez anciens en Belgique du fait de la configuration géoadministrative du pays. Il s'agissait sûrement d'une omission chez Larousse. En outre, on remarquera entre *région*, *régionalisation* et *régionaliser* trois manières de marquer la particularité belge : soit par le cumul (« En France et en Belgique »)²¹, soit par la désignation spatiale (« En Belgique »), soit enfin en précisant le territoire national d'emploi (« En Belgique »). Cette diversité de la syntagmatique des marques diatopiques est aussi relevée dans le *GR* et le *PR* sans qu'une explication métalinguistique précise figure dans les préfaces correspondantes. D'ailleurs les marques topolectales ne sont pas indiquées dans la table des matières comme le signale Nathalie Gasiglia dans le présent volume.

A cet égard, Alain Rey (1990 : 25)²² fait remarquer des ambiguïtés du même type relatives à la marque de domaine tantôt nom d'agent, tantôt phrase verbale métalinguistique ou passage entre domaine thématique et caractérisation d'une langue de spécialité, etc. Cela pourrait correspondre, il nous semble, à une absence de systématisation de la technique lexicographique en partie due à l'absence d'outils informatiques puisque l'importance quantitative des données rend difficile la mémorisation des habitudes syntaxiques et des formules lexicographiques²³.

20 Il serait ici intéressant de souligner qu'en 2003 et 2004 le groupe de recherche Valibel a bien fait d'élargir son corpus de dictionnaires français en nous demandant de prendre en compte le *Trésor de la langue française* et le *GR* au lieu de nous limiter à la dernière édition des DGM parus. En effet, l'absence de vue diachronique sur la nomenclature française peut entraîner des erreurs d'évaluation quant au caractère belge d'un item, car il a pu, au cours des années, être supprimé du dictionnaire.

21 On pourrait ici évoquer ce que Galisson nomme la « charge culturelle partagée ».

22 Sur le manque d'harmonisation des marques topolectales dans le *PR*, voir Poirier 2003 : 208. Notons que cet article est particulièrement exemplaire (v. note 15) pour illustrer la position particulièrement ambiguë de Poirier vis-à-vis du français ou du français hexagonal, tantôt admiratif, tantôt critique, cela rappelant évidemment l'épisode de cessation des enseignements d'été de Georges Straka au Québec à la fin des années 1970.

23 Notons en outre que la précision sur la situation de discours côtoie la caractérisation des lieux géographiques d'usage : « terme de palais » ou « se dit au palais » chez Furetière (1990 : 26). Le chevauchement des marques stylistique et diatopique s'explique par le fait que les usages sont encore souvent connotés au début du XX^e siècle par la situation sociolinguistique. Ainsi la marque *littéraire*, fort critiquable puisque le discours littéraire n'est pas spécifié, pourrait parfois être assimilé à la vision du français parisien selon la perception du français au Canada d'Adjutor Rivard, opposée au français régional franco-canadien (Verrault, 2006 : 36 et Rey, 1990 : 17).

L'hétérogénéité des indications géographiques constatée dans la production lexicographique depuis le XVII^e siècle nous conduit à distinguer deux types d'indications :

1. La caractérisation des lieux d'usage par la grammaticalisation de la métalangue lexicographique (*marque diatopique macrostructurelle*). Prenons comme exemples « Se dit en Belgique », « Dans les provinces du Sud on disait autrefois... ».
2. L'indication topolectale (après l'entrée) par une abréviation ou un symbole typographique (chez Louis-Alexandre Bélisle, Richelet et Marie-Eva de Villers par exemple) (*marque diatopique microstructurelle*) (Bray, 1990 : 44). Il s'agit de marques comme « Belgique », « Canada (Québec) » ou « Nord-Pas-de-Calais ».

CONCLUSION

Au terme de cette réflexion sur les topolectismes et les marques diatopiques, il convient de faire remarquer – malgré les efforts déployés depuis quelques années par la Belgique pour décrire les particularités de ses variétés de français – la sous-représentation des belgicisms relevant de l'administration et de la politique belge comme *bi-caméralisme* (également attesté au Québec), *numéro privé* (qui ne semble pas avoir d'équivalent en France (*liste rouge*) et au Québec (*liste noire*)), *percepteur* (dans le sens postal du terme), *Centre de guidance* (comme syntagme²⁴), etc. Notons aussi les particularités phonétiques de certains toponymes belges comme *Anvers* prononcé à Paris sans [s] final ou *Bruxelles* prononcé à la française comme cela s'écrit et non à la manière plutôt flamande qu'ont les Belges de prononcer le nom de leur capitale²⁵. Par ailleurs, si le vocabulaire estudiantin a fortement été pris en compte, notamment à cause de l'influence qu'exerce la population estudiantine néolouvainiste (de Louvain-la-Neuve) sur les dictionnaristes, il me semble que l'on n'a pas encore exploité toutes les richesses de mots comme *décuver*, *kicher*, *se prendre une guinze* ou de *jobbiste* que l'on pourrait mettre en comparaison avec *jobbeur* attesté au Québec ou enfin *journal de classe* équivalent à *cahier de textes* (France) et *cahier de devoirs* (Québec).

D'éventuelles contributions pourraient explorer la dimension sémantique des topolectismes, par exemple le fait qu'un topolecte puisse se retrouver dans deux cas de figures équivalent plus ou moins synonyme d'un lexème existant déjà dans le français de référence ou sans équivalent. Cette pratique nécessiterait également sur le plan lexicographique de créer des renvois pertinents vers ces mots. En somme l'équivalence topolectale est-elle plus ignorée que la particularité lexicale et référentielle sans équivalence dans le français de référence ? C'est là, il me semble, une question que peu ont abordée jusqu'ici.

Au fil des siècles, la technique lexicographique s'affinant, les indications diatopiques se sont unifiées. Ainsi, Francine Mazière affirme que la marque en tête se réduit au nom de domaine ou à la notation du contexte d'énonciation. « On voit tout ce qu'elle a ainsi perdu du rapport langue / énonciateurs / société.

24 Lors du colloque qui a donné lieu à cette contribution, Lola Devolder m'a fait remarquer que *centre de guidance* est attesté en France, contrairement aux sources dictionnaires que nous avons consultées.

25 Remarque d'Henriette Walter.

Et il n'est pas certain qu'elle ait clarifié le problème » (Mazière 1990 : 101). C'est l'aspect réducteur et forcément elliptique de l'indication diatopique qui sur le plan dictionnaire relève d'un choix difficile et révélateur du caractère forcément optionnel de tout dictionnaire de langue. Encore faudrait-il nuancer cette affirmation en précisant que les topoclectismes sont variablement marqués d'un dictionnaire à un autre selon le projet dans lequel ils s'inscrivent. C'est là un autre sujet qui mériterait d'être abordé dans des contributions futures, faute de quoi on ne pourra prétendre à avoir parfaitement *contourné la marque*.

En s'inspirant de Jean-Claude Boulanger, on pourrait se demander à quels critères répond l'inclusion de nouveaux topoclectismes dans un DGM. Les grands dictionnaires monolingues de chez Robert, Hachette et Larousse procèdent souvent à l'inclusion de néonymes, de lexèmes ou d'unités sur la base de leur fréquence. Mais on peut difficilement expliquer pourquoi un dictionnaire important comme le *GR* est resté plus de vingt ans sans ajouter de belgicismes avant que nous en fassions part à l'éditeur en 2003. Les écarts entre les données provenant du terrain et celles de la nomenclature de l'inventaire ne permettent pas de constater un dynamisme certain sur le plan variationnel, puisque les DVF, comme ceux qui ont été récemment publiés au Québec par l'équipe de Claude Poirier, en Suisse par Pierre Knecht et André Thibault, en France par Rézeau et prochainement en Belgique par l'équipe de Michel Francard, ne sont pas soumis aux mêmes contraintes éditoriales que les DGM. Le *Dictionnaire historique du français québécois*, du fait qu'il est inachevé, répond plutôt au critère de l'exhaustivité de traitement de certaines unités alors que les deux autres ont tenté de décrire le plus largement possible leurs usages nationaux en faisant état également des topoclectismes de partage.

RÉFÉRENCES BIBLIOGRAPHIQUES

- DE BALZAC Honoré, 1842, *Une double famille*, Paris, Gallimard, 1976.
- DE BALZAC Honoré, 1843, *Le Père Goriot*, Paris, Gallimard, 1976.
- BÉJOINT Henri, 1981a, « Les marques d'usage dans les dictionnaires de langue anglaise », *Trema*, 7, p. 69-74.
- BÉJOINT Henri, 1981b, « Variété de langue et marques d'usage dans les dictionnaires », *Grazer linguistische Studien*, 15, p. 7-16.
- BOULANGER Jean-Claude, 2001, « L'aménagement des marques d'usage technoclectales dans les dictionnaires généraux bilingues », dans Jean Pruvost (éd.), *Les Dictionnaires de langue française. Dictionnaire d'apprentissage, dictionnaires spécialisés de la langue, dictionnaires de spécialité*, Paris, Champion, p. 247-271.
- BRAY Laurent, 1990, « Les marques d'usage dans le *Dictionnaire françois* (1680) de César-Pierre Richelet », dans M. Glatigny (éd.), *Lexique*, 9, *Les Marques d'usage dans les dictionnaires (XVII^e-XVIII^e siècles)*, p. 43-59.
- CHARLIER Gustave et HANSE Joseph, *Histoire illustrée des lettres françaises de Belgique*, Bruxelles, La Renaissance du livre, 1958.
- CORBIN Pierre, 1989, « Les marques stylistiques / diastratiques dans le dictionnaire monolingue », in F. J. Hausmann, O. Reichmann, E. Wiegand und L. Zgusta (Hrsg.), *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicogra-*

phy / Encyclopédie internationale de lexicographie, 1, Berlin u. New York, De Gruyter, p. 673-680.

- DE SURMONT Jean Nicolas, 2004, « Caractéristiques et variations du marquage diatopique », in G. Williams and S. Vessier (eds.), *Proceedings of the 11th Euralex International Congress, Euralex 2004, Lorient, July 6-10, 2004*, vol. 3, Lorient, Université de Bretagne Sud, p. 779-786.
- DE SURMONT Jean Nicolas, 2005, « Du contexte à la citation : les récents développements de la dictionnaire », dans D. Blampain, P. Thoiron et M. Van Campenhoudt (éds), *Mots, termes et contextes. Actes des septièmes Journées scientifiques du réseau de chercheurs « Lexicologie, terminologie, traduction »*, ISTI, Bruxelles, 8-10 septembre 2005, Paris, Éditions des archives contemporaines ; <http://www.termisti.refer.org/colloque/journltt02.htm>
- DUBOIS Jean et DUBOIS Claude, 1971, *Introduction à la lexicographie: le dictionnaire*, Paris, Larousse.
- GALISSON Robert, 1999, « La pragmatique lexicoculturelle pour accéder autrement à une autre culture par une autre culture, par un autre lexique », *Études de linguistique appliquée*, 116, p. 477-496.
- GARNIER Yves, 2005, « Les francophonismes dans le *Petit Larousse* », dans M. Cormier et A. Francœur (éds), *Les Dictionnaires Larousse, genèse et évolution*, Montréal, Presses de l'Université de Montréal, p. 219-248.
- GIRARDIN Chantal, 1987, « Système des marques et connotations sociales dans quelques dictionnaires culturels français », *Lexicographica*, 3, p. 76-102.
- GLATIGNY Michel, 1998, *Les Marques d'usage dans les dictionnaires français monolingues du XIX^e siècle : jugements portés sur un échantillon de mots et d'emplois par les principaux lexicographes*, Tübingen, Niemeyer.
- HANSE Joseph, 1987, *Nouveau Dictionnaire des difficultés du français moderne*, 2^e éd., Paris et Gembloux, Duculot.
- HANSE Joseph, 1994, *Nouveau Dictionnaire des difficultés du français moderne*, Louvain-la-Neuve, De Boeck-Duculot, 3^e éd. établie d'après les notes de l'auteur avec la collaboration scientifique de D. Blampain, avant-propos de Ghislaine Hanse et D. Blampain.
- HANSE Joseph et BLAMPAIN Daniel, 2000, *Nouveau Dictionnaire des difficultés du français moderne*, 4^e éd., Bruxelles, De Boeck-Duculot.
- KLINKENBERG Jean-Marie, 1998 et 2002, Notes inédites accompagnant les recommandations faites à Larousse pour l'ajout ou la suppression de belgicismes.
- LARA Luis Fernando, 2004, *Lengua histórica y normatividad*, Mexico, El Colegio de México.
- LEBOUC Georges, 2006, *Dictionnaire de belgicismes*, Bruxelles, Éditions Racine.
- MERCIER Louis et VERREAULT Claude (éds), 1998, *Les Marques lexicographiques en contexte québécois. Actes de la table ronde tenue à Montréal les 3 et 4 novembre 1994*, Québec, Gouvernement du Québec, Office de la langue française (Études, recherches et documentations).
- MAZIÈRE Francine, 1990, « Les marques de fabrique. Marquage et marques de domaine dans les dictionnaires classiques, du Furetière aux Trévoux », dans M. Glatigny (éd.), *Lexique*, 9, *Les Marques d'usage dans les dictionnaires (XVII^e-XVIII^e siècles)*, p. 89-111.

- MCARTHUR Tom, 1986, *Worlds of reference : Lexicography, Learning and Language from the Clay tablet to the Computer*, Cambridge, Cambridge University Press.
- MESSELAAR Petrus Adrianus, 1988, « Les marques “familier” et “populaire” envisagées des points de vue lexicologique et lexicographique », *Cahiers de lexicologie*, 53 (2), p. 91-106.
- PIRON Maurice, 1973, « Les belgicisms lexicaux : essai d'un inventaire », *Travaux de Linguistique et de Littérature*, 11(1) (*Mélanges de linguistique française et de philosophie et littérature médiévales offerts à M. Paul Imbs*), p. 295-304.
- POIRIER Claude, 2003, « Variation du français en francophonie et cohérence de la description lexicographique », dans M. Cormier, A. Francœur et J.-C. Boulanger (éds), *Les Dictionnaires Le Robert*, Montréal, Les Presses de l'Université de Montréal, p. 189-226.
- PRUVOST Jean, 2005, « Quelques concepts lexicographiques opératoires à promouvoir au seuil du XXI^e siècle », *Études de linguistique appliquée*, 137, « Dictionnaires et innovation dans les dictionnaires français », p. 7-37.
- PRUVOST Jean, 2006, *Les Dictionnaires français, outils d'une langue et d'une culture*, Paris, Ophrys.
- REY Alain, 1990, « Les marques d'usage et leur mise en place dans les dictionnaires du XVII^e siècle : le cas Furetière », *Lexique*, 9, p. 17-29.
- DE ROBILLARD Didier, 1993, « Normalisation de la régionalité, régionalisation de la norme », dans Daniel Baggioni (éd.), *Encyclopédies et dictionnaires français (Problèmes de norme(s) et de nomenclature)*, Aix-en-Provence, Publications de l'Université de Provence, p. 141-173.
- VALIBEL, 2004, *Les Particularités lexicales du français en Belgique francophone. Participation du Centre de recherche Valibel au projet du Dictionnaire Universel Hachette*, tapuscrit.
- VERREAULT Claude, 2006, « Conception du français au Canada selon Adjutor Rivard, principal artisan de la Société du parler français au Canada », dans C. Verrault, L. Mercier et T. Lavoie, 1902-2002, *La Société du parler français au Canada, cent ans après sa fondation : mise en valeur d'un patrimoine culturel*, Québec, Les Presses de l'Université Laval, p. 29-54.

SIMPLY TOO MANY THINGS TO LABEL! DEALING WITH THE LEXICOGRAPHICAL NEEDS OF BIDIALECTAL SPEECH COMMUNITIES

par Pavlos PAVLOU
Université de Nicosie

Résumé — Les dictionnaires sont écrits en ayant comme point de référence les langues et non pas les communautés linguistiques. Cependant, étant donné que les variantes régionales n'y sont pas souvent incluses, cette pratique ne répond pas aux besoins des communautés linguistiques faisant usage d'une variété non standard, qu'elle soit utilisée de manière exclusive ou en parallèle avec la variété standard. En dehors la variation diatopique, d'autres types de variation liées surtout à l'appartenance sociale et à l'âge qui diffèrent des variétés semblables présentes dans la communauté standard, sont aussi ignorés par les dictionnaires. Dans cette communication nous expliquons de quelle manière les dictionnaires bilingues du grec manquent de répondre aux besoins des locuteurs, aussi bien natifs que non natifs, de la communauté chypriote grecque.

Mots clefs — communauté langagière, faux-amis, langue minoritaire, langues pluricentriques

Abstract — Dictionaries are written making reference to languages and not to linguistic communities. However, given that regional variations are not often included in lexicographical data, this practice does not answer needs of linguistic communities using non standard varieties. This paper discusses this topic with the example of English loanwords in contemporary Cypriot Greek. By means of the results on a survey focusing on the relationship between English loanword knowledge and use and age, the paper discusses the potential position of such words in dictionaries of Cypriot-Greek. The survey indicates that there is an age-related stratification of these loanwords and thus their marking in the dictionary becomes problematic. Another finding of the study is that some loanwords have different meanings for male and female members of the speech community regardless of age. Again, the extent to which this information should also be included in a dictionary is also debated. Finally, it is investigated whether loanwords have the same meaning and if they receive the same grammatical gender in dictionaries of Standard Modern Greek and Cypriot Greek.

Keywords — bidialectalism, speech community, false friends, pluricentric languages.

INTRODUCTION

Contrary to what most people believe the majority of speech communities in the world are bilingual / multilingual or bidialectal / multidialectal. In the case of bidialectal or multidialectal societies, one dialect serves as the standard dialect. The reasons that confer this status to a given linguistic variety are extralinguistic ones and usually relate to the relative power the users of this community have, be it economic, military, political, educational or literary. More often than not, this bidialectalism exists within one state as was the case in Greece up to 1976 where two varieties, *Katharevousa* (puristic language) and *Demotic* (popular language) were used, a situation described by Ferguson (1959) using the term *diglossia*. A similar example is a speech community on Sicily in which standard Italian is used along with the local variety.

However, the phenomenon of bidialectalism, becomes more complicated in cases where the standard variety is, at the same time, the standard language of another state. For example the Austrian speech community employs both a number of local dialects along with High German which is the language of Germany. The same holds for Cyprus, where the members of the Cypriot-Greek speech community use both the sub-varieties of Cypriot Greek (CG) and Standard Modern Greek (SMG) to various degrees of competency. At the same time, SMG is the official language of another state, Greece. In such situations, that is when two or more states share the same linguistic code we can talk of *pluricentric languages* (Clyne 1992). In some settings, a pluricentric language has the same status and develops equally in the two states. We refer to such cases as *symmetric* case of a pluricentric language, because no speech community (and consequently the linguistic variety employed by this community) clearly dominates culturally. An example of this would be English which develops more or less equally at all centers it exists (USA, UK, Canada, Australia, etc). However, there are many cases in which one speech community (and its linguistic variety) is considered dominant, either because of the sheer number of the speakers or in combination of other factors such as economy or cultural power. This is termed as an *asymmetric* case of a pluricentric language. Standard German in relationship to varieties could be considered an example of this.

The linguistic situation in Cyprus is another, clearly recognizable example of an asymmetric pluricentric language. In my opinion, another factor in the definition of a symmetric or an asymmetric pluricentric language is the language attitudes of the speakers of the two varieties. In many cases, the speakers of the dominant variety do not understand or use the non-dominant variety whereas the speakers of the non-dominant variety are bidialectal. If, as is the case in Cyprus and many bidialectal communities, they do not hold such positive attitudes towards their variety, this variety will not be further cultivated, it will wither and eventually disappear. According to sociolinguistic research, the CG and its use are not fully appreciated (see Papapavlou and Pavlou 1998, Pavlou 1999, among others). Such conditions discourage and prevent the standardization of the linguistic code and as a result members of the speech community rely on reference books (grammars, dictionaries, etc.) of the standard variety for their needs. Such a need would be the use of bilingual dictionaries when members of the Greek-Cypriot speech community learn a foreign language or when they are involved

in translations. Moreover, such bilingual dictionaries are needed for non-native speakers of Greek who reside on the island and need to communicate in Greek.

However, the existing SMG based dictionaries do not completely serve the needs of the members of the Greek-Cypriot speech community, be it autochthonous or not, since the information they contain pertains to the linguistic code, society and culture of the dominant speech community, namely SMG. This paper examines the ways in which SMG based bilingual dictionaries are inadequate for the members of the Greek-Cypriot speech community. Of course, there is also another dimension to the absence of reference materials for CG, namely the assumption that fostering the Cypriot Dialect would remove Cypriot Greeks away from the hellenic ethnos. Barbato and Varvaro (2004: 432) in their discussion on dialect dictionaries in Italy, state that “defence of the dialect often appears in conservative light and promotes a separatist identity”. However, this issue is beyond the scope of the current paper.

Besides the need for a dialect dictionary mentioned above, Barbato and Varvaro (2004) list a few additional purposes and justifications such as purely scientific ends, the assistance they can offer to the dialectal speaker to learn the standard variety, an effort to show the dialect’s wealth, nobility and remote origins.

THE SOCIOLINGUISTIC SITUATION IN CYPRUS

Cypriot Greek is a Modern Greek Dialect, which belongs to the South-Eastern group of modern Greek dialects and bears many similarities with the dialects spoken in the Dodecanese Islands. Both structural and sociolinguistic aspects of the dialect have been studied and there are numerous studies focusing on how language and society interact on Cyprus. These include among others Siricha (1995), Papapavlou & Pavlou (1998). Cyprus is, in some respects, a speech community which presents certain idiosyncrasies from both the linguistic and sociolinguistic point of view. CG consists of a variety of regional varieties and registers which constitute a dialectal continuum. CG speakers make active use, in various degrees of competency, of many segments of the Cypriot dialectal continuum. This continuum consists of various levels of both the so called Cypriot Greek Variety and Standard Modern Greek (see Karyolemou 1999, Tsiplakou *et al.* 2006). Even though it is not the standard dialect, it is still robust, has a large number of speakers and is spoken in a geographic area which has the status of a state, the Republic of Cyprus.

One rather unique feature is that Cyprus is officially a bilingual state in Greek and Turkish, with very few Greek Cypriot speakers being competent in Turkish, whereas many Turkish Cypriots, especially those belonging to older generations are quite fluent in mostly mesolectal and acrolectal varieties of the Cypriot dialect continuum.

Moreover, English plays a significant role in the linguistic repertoire of Greek Cypriots as Cyprus is an ex British colony (see Davy & Pavlou 2001, McEntee & Pouyioukas 1997, Papapavlou 1997). The composition of the CG speech community is depicted in Table 1:

Ethnic and religious groups	Linguistic repertoire
Greek Cypriots	Greek Cypriot Dialect, Standard Modern Greek, English
Turkish Cypriots	Turkish Cypriot Dialect, Standard Turkish, Greek Cypriot Dialect, English
Maronites	Greek Cypriot Dialect, Cypriot Arabic (endangered language), English
Armenians	Greek Cypriot Dialect, Armenian, English
Latins	Cypriot-Greek, English, (Italian, French)
Foreign residents	Russian, Georgian, Serbian, Tamil, Arabic, Polish, Rumanian, Bulgarian, Pakistani, Chinese, Tagalog, Greek Cypriot Dialect, Standard Modern Greek, English

Table 1.- The Greek-Cypriot speech community

DICTIONARIES AND LANGUAGE USERS

A bilingual dictionary usually refers to languages and not speech communities. We invariably refer to a “French-Greek dictionary”, an “English-Greek dictionary”, etc. A dictionary is a reference work used by a native speaker of a language who is looking for the equivalent of a word in his/her first language, in another language. Moreover, it is used by learners of a second or a foreign language. Second language learning describes a situation where the learner learns a language other than his/her mother tongue in a speech community whose members make use of this language. It may be the only language used in this speech community or there may be other languages or varieties used in this speech community. In the latter case, in order for the setting to be considered as a “second language” setting, the numbers of native or non-native speakers of the linguistic code to be learned must ensure frequent and adequate opportunities for communication. For example, English in Cyprus is, to some extent and under certain definitions, a second language. That is, the setting allows for a frequent and varied interaction and communication by means of English (Davy & Pavlou 2001).

A learner of Greek as a second language in Cyprus learns Greek to communicate with the native speakers of the Cypriot Dialect and other non-native speakers. In that case CG is enough. In cases where non-native speakers of the Cypriot Dialect need to communicate with mono-dialectal speakers of SMG, then they may need to be in command of SMG.

In order for a dictionary to be useful to the members of a specific speech community it should reflect the linguistic repertoire of the members of this speech community. Sometimes the speech communities make use of the language described in the dictionary. This is the case of speech communities which make use of the standard dialect. Actually, it is not the case that they make use of the standard dialect, rather, the members of this speech community have been able to elevate their dialect to the status of the standard variety.

If a dictionary is based on the standard, then the dialectal speaker may not be completely served by this dictionary and, therefore, s/he would need a dictionary which includes dialectal forms. Still, the dialectal speaker's need for such a dictionary may be questioned by some. This language user is to some degree bidialectal and by definition and s/he is in command, to a certain extent, of the standard which s/he has learned mostly through education and the media which are the two major domains which the standard dominates in diglossic or bidialectal situations. This is certainly true for the Cypriot setting.

This may be true, to some extent, for communities who are not only linguistically but also economically peripheral and depend on the linguistic centre which is often the economic centre as well. However, Cyprus is not like that. It is a prosperous society which attracts foreign workers who need to communicate with Cypriots mostly through the local variety the CG.

There are around 110 000 immigrants in Cyprus and for many of them English is a *Lingua Franca* e.g. Sri Lankan women. These women have no right to venture into news fields of entrepreneurship, speaking English is virtually a requirement for their job and therefore English is enough for them to survive in Cyprus.

However, immigrants of European origin need and are very keen on learning Greek. Many of them have the right to stay as long as they want and many do not speak English. Therefore, if they want to live and work on Cyprus, familiarity with the CG variety is necessary. Upon arriving to Cyprus, they become aware of the need to learn Greek and of the sociolinguistic position of the varieties spoken in Cyprus since they are familiar with similar situations in their countries. They have a dilemma as to what to learn. Pavlou and Christodoulou (2001), discuss various issues relating to teaching and learning Greek as a second language in Cyprus. One particular issue is reference books on CG such as grammars and dictionaries. Dictionaries of the Cypriot Dialect would be of help to them.

There are quite a few dictionaries of CG (see Katsoyannou 2006) but they do not seem to be of any real usefulness to non native speakers of Greek residing and working on Cyprus for a number of reasons. First, they resemble "glossaries" i.e. they are a compilation of the lexical items that differentiate CG from SMG. Moreover, they constitute a description of CG with reference to SMG, i.e., they describe the forms that deviate from the SMG and elements that are present in CG but do not exist in SMG. They start with the (erroneous) assumption that the SMG is the correct form of the language and that CG is a deviation of the standard. Such dictionaries are of interest to linguists and not to people who would like to learn the language in order to communicate with other members of the speech community. They are more like museum exhibits since they contain many obsolete words which are there in order to be observed and admired but not to be used. Non-native speakers who need Greek in order to communicate with native speakers of CG must be in command of CG, which besides the lexicon that it shares with SMG, possesses a large number of words that are to be found only in the dialect.

The problems mentioned above along with other related issues are studied by the branch of lexicography called dialectal lexicography (see Weiss 2003, Giandomenico 2006, Kagaine & Timuska 1993).

INADEQUACIES OF SMG ORIENTED DICTIONARIES FOR THE CG USER

The study of lexicography is a recent field in Cyprus and there are not so many dictionaries of the dialect (see Katsoyannou 2006). This section focuses on the numerous ways existing dictionaries of Greek fail to satisfy the needs of the Greek-Cypriot user because of differences between the two codes both on the structural level (phonology, morphology, syntax, semantics) and on the socio-pragmatic and cultural level. For the major differences between the two codes, see Pavlou and Christodoulou (2001), Coutsougera *et al.* (2008) among others.

STRUCTURAL DIFFERENCES

Phonological information

Phonology is the linguistic level which distinguishes unmistakably CG from the rest of the Modern Greek Dialects including of course SMG (see Karyolemou & Pavlou 2001). The Cypriot dialect speaker is immediately recognised by speakers of other Greek dialects and as far as mutual intelligibility is concerned the speakers of CG understand to a large extent speakers of other dialects, mainly because of their exposure to the media which invariably adopt the SMG (see Pavlou 2004). On the other hand, SMG speakers or speakers of other Greek dialects have difficulty in understanding a Cypriot speaker especially if they are not speakers of a related dialect (South-Eastern group) and have limited exposure to CG speakers. Even though CG has such a distinct phonology the distinctive way words are pronounced in CG is not marked in a bilingual dictionary of Greek. This could help the user who is not competent in SMG to use the dictionary appropriately. The inclusion of phonological information can only be achieved after codification and standardization, and the development of a linguistically sound transcription system of CG. This is necessary since certain distinctive sounds of CG can not be rendered by means of the existing letters and diacritics of SMG.

Morphological information

There is a number of morphological differences between the two dialects, which need to be mentioned in a bilingual dictionary of Greek that would be useful both for the native and the non-native speaker of CG. For example, if a native speaker of CG wants to find the equivalent of *της* (pronoun, feminine, accusative, plural) in French, s/he may not find it in a Greek-French dictionary because the entry for this word in the dictionary reflects SMG morphology and phonology. Therefore, the entry would be *τις*. As stated earlier the native speaker of CG is bidialectal and it is highly likely that s/he is aware of the SMG conventions but still it is not something that should be taken for granted.

The absence of morphological information on CG in the dictionaries we are discussing may cause difficulties for the non-native speaker of CG. For example, the non-native speaker may hear the word *κάνουν* (they do) and attempt to find its meaning in a dictionary. The non-native language learner will not be able to locate the word since it appears as *κάνουν*, the SMG form, which is quite different from the CG one.

Moreover, existing dictionaries do not indicate the different gender, and the article that precedes nouns in CG. For example the word for “butter” in CG is both masculine and neutral (*ο βούτυρος, το βούτυρο*). Some other common words in this category are:

<i>Sugar</i>	<i>Ο</i> (masc.) / <i>η</i> (femin.) <i>ζάχαρης -ις, το</i> (neutral) <i>ζάχαρι</i>
<i>Sand</i>	<i>η</i> (femin.) <i>άμμος, ο</i> (femin.) <i>άμμος</i>
<i>Paradise</i>	<i>ο</i> (masc.) <i>παράδεισος, η</i> (femin.) <i>παράδεισος</i>

Syntactic information

There are few differences between CG and SMG, the most notorious one being personal pronoun proclisis and enclisis alternations or clitic-second phenomena. For example the phrase “I told him” appears both as SMG *Του Είπα* (*him said-I*) and CG *Είπα του* (*said-I him*) with the latter being much more frequent in the spoken language. The non-native speaker will never know whether this is free variation or a cue to different meaning. Therefore, this information could be included in a dictionary as well.

LEXICON

False friends

SMG based dictionaries do not give information on false friends i.e. words in the two varieties that have the same phonological and morphological form but have a different meaning (see Szpila 2006). There are a number of such words and they can be classified in various classes (see Arvaniti 2002, Hatzopoulou 2006). Some of them are quite frequent in everyday speech such as the word *ανηψιός* (nephew in SMG and cousin in CG), or *βαλίτσα* (luggage in SMG, luggage and school bag in CG).

When resorting to, for example, a French-Greek dictionary for the word *cousin* the learner will produce a translation *εξάδελφος* which may not be understood by all, especially, older dialectal speakers. When using a Greek-French dictionary in order to locate the equivalent of *ανηψιός* in French, the learner will come up with “neveu” (nephew), which is not of course the real meaning of this word in the Cypriot setting.

Words used only in Cyprus

Dictionaries of SMG do not contain words for concepts that exist both in the CG and SMG speech communities and for which the two codes have different words and not just different ways of pronouncing the same word. Such words are *χρυσόμηλο* (apricot), *κούσπος* (axe), *τσαέρα* (chair). These three words are chosen because they are sociolinguistically, stylistically or otherwise (e.g. age related) stratified. *Χρυσόμηλο* (apricot) is a word that would be used by all speakers regardless of the level(s) of the CG dialect continuum they are employing. One reason for this would be the fact that, even though it does not exist in SMG, this word does not violate the phonotactics of SMG and therefore it is not sociolinguistically marked. The second and third words are more sociolinguistically of stylistically stratified since people may use them or not depending on the default level or targeted level of the dialect they are using. Again, the two words do comply with the SMG phonological rules and processes of the SMG. On the other hand, there are words like *πατίτσα* (watermelon) which violate SMG pho-

nology and therefore are sociolinguistically of stylistically stratified and stigmatised. In the latter case, this should be mentioned in a dictionary. In other words, sociolinguistic information should be included in dictionaries of Greek.

Pragmatic differences

It is widely known that words, besides their denotational or referential meaning also possess social meaning or connotation. Social meaning is the meaning a word gets in the social context it is used. This information could be termed “pragmatic information”. Current dictionaries do not give the denotational meaning of words that have different pragmatic meaning in Cyprus. This has been observed widely for many languages (see Wen 2007).

Therefore, besides semantic differences between the lexicons of the two dialects there are differences in the area of pragmatics i.e. in the way the same words are used in the two speech communities. Bidialectal speakers, with either dialect being their dominant one, recognize that certain words in the two codes are not used in the same way. For example *σπίτι* in SMG means “home, domicile” and includes “apartments”, whereas in the CG speech community it means “home, domicile” but only detached or semi-detached houses and not apartments.

Another example is the use of the word *πρόνοιες* which in general means provisions in the two dialects. However, for “the provisions of the law”, “provisions of the plan” and other similar expressions, the members of the SMG speech community use the word *διατάξεις*. The use of *πρόνοιες* by the native speakers of CG for these cases as well, could be the result of translations of English words or phrases.

Another example of a pragmatically based difference is the following. A common meaning, especially among younger generations of the word *Κινέζος* (Chinese) is OMONOIA football team fan, the connection being that both groups have a very large number of members. When a speaker makes use of the non-literal meaning of this word, the interlocutor, who may not be aware of the metaphorical meaning the word has in the Cypriot setting, may not understand why the Chinese are spoken about. Of course one could claim that this kind of information should be omitted in dictionaries because they are slang and their popularity will, sooner or later, fade out.

Cultural differences

According to the Whorf-Sapir Hypothesis culture is reflected in language. Even though at a macro-level the two speech communities share the same culture, they have certain cultural differences as well which are of course reflected in the language of the two communities. These cultural differences are mainly due to geographical isolation, the different cultural influences and the existence of two different states. Therefore, there are words representing manifestations of the particular cultures or state institutions, that are absent from either communities. For example, the famous Cypriot cheese *χαλλούμι* is not produced in Greece and therefore is not included in bilingual Greek dictionaries, even though it can be, rather seldom, found in monolingual Greek dictionaries (e.g. Babinotis 1998). There are some institutional terms that exist only in one of the two communities or are not quite the same in the two communities. For example, in SMG we have

two terms for law enforcement personnel. One is *αστυνομικός* (policeman) which is a concept existing in both communities and expressed through the same lexical item. However in Greece there is the institution of *χωροφύλακας* which does not exist in Cyprus. Also, the terms *ευρύτερο δημόσιο* (wider public sector) and the distinction between I.X. (for private use) and Δ.X. (for public use) with reference to cars, do not exist in CG, even though the concepts do exist. A similar case is the word for ombudsman, a concept which is rendered through *Επίτροπος Διοικήσεως* in Cyprus and *συνήγορος του πολίτη* in Greece. A user of CG seeking the translation of *επίτροπος Διοικήσεως* in any bilingual dictionary of Greek will not find an entry for this. The native speakers of CG are not aware of the fact that this word does not exist in SMG.

Sub-regional and age-related varieties

Next to dialectal speakers we also have speakers of sub-dialects who also have linguistic rights such as the right to the standardization and documentation of their variety. According to Newton (1972), CG has seventeen such sub-varieties and the question rises whether these varieties should be included in a dictionary of SMG. On the theoretical level they should be, but there are two problems associated with the adoption of such a policy. First, these sub-dialects must be adequately described and second, the volume of the dictionary should also be a consideration. Unabridged dictionaries take time and money to compile, are voluminous and not easy to use for the foreign language learner.

Another issue which will be briefly broached is whether age-related variation with reference to both the indigenous lexicon and borrowings should be included in dictionaries. Pavlou (2007) showed that there is both age-related stratification and semantic differentiation according to gender, of borrowings in CG. Two arguments against including such words in a general bilingual dictionary would be the fact that some of these words are “trendy” and will soon disappear and that such marking would constitute excessive information that would render dictionaries enormous and difficult to use.

CONCLUSION

This paper has discussed the issue of regionalisms, that is linguistic items that belong to dialects other than the standard, in lexicography. Ideally, bilingual dictionaries should include words – and the necessary annotation on them – from varieties other than the standard. In the case of Greek this would increase the volume of a dictionary substantially since, at least CG, not to mention other Greek dialects, differ substantially from SMG in all linguistic levels including lexicon. Moreover, CG serves as an instrument for communication in a robust and thriving speech community. In contrast to other more vulnerable Modern Greek Dialects (such as Pontic Greek) which are more likely to be lost because of the prominence of the standard, CG is still strong. However, CG faces threats from SMG but also from English, the same way most world languages do.

A decision to include lexicographical information from regional dialects should affect not only CG but all the Modern Greek Dialects regardless of their capacity. Therefore, dialectal information in dictionaries could refer to groups of dialects. For example, CG is one of the South-Eastern dialects of Greek and therefore shares features with them. Therefore, a word that exists in all these

dialects such as *ποσκόλιο* “passe-temps” and not in the standard could be marked as belonging to dialects of Cyprus, Karpathos, Rhodes, Kastellorizo and Hios, or simply as “South-Eastern dialects”. This is an economical solution but it requires some general dialectological knowledge.

Moreover, I would like to add that dialectal dictionaries or dictionaries that cover dialectal varieties substantially, are important from both the linguistic and cultural point of view. We have seen that such dictionaries can reflect the exact linguistic variations and semantic nuances. Also, they will reflect the idiosyncrasies and distinct cultural traits of a specific speech community.

Finally it should be mentioned that CG differs substantially from SMG and therefore learners of Greek as a foreign language need to learn CG in order to function fully in the Greek-Cypriot speech community. A dictionary of CG would be helpful to them. In addition, such dictionaries and other similar reference works will be instrumental in the effort to standardising the dialect and therefore helping it survive in relationship to SMG in the homogenizing climate of globalization.

REFERENCES

- ARVANITI Amalia, 2006, “Linguistic practices in Cyprus and the emergence of Cypriot Standard Greek”, *San Diego Linguistic Papers*, 2, p. 1-25.
- BABINIOTIS Georgios, 1998, *Dictionary of Modern Greek*, Athens, Lexicology Center.
- BARBATO Marcello and VARVARO Alberto, 2004, “Dialect dictionaries”, *International Journal of Lexicography*, 17 (4), p. 429-239.
- CHATZOPOULOU Konstantinia, 2011, “False friends in Modern and Cypriot Greek”, in Marilena Karyolemou, Fabienne Baider et Marianne Katsoyanou (éds), *XXX^e Colloque international de linguistique fonctionnelle*, Limoges, Lambert-Lucas, p. 37-40.
- CLYNE Michael (ed.), 1992, *Pluricentric Languages: Differing Norms in Different Nations*, Berlin, Mouton de Gruyter.
- DAVY Jim and PAVLOU Pavlos, 2001, “Is Cyprus an ESL country?”, *Proceedings of the 12th International Conference of the Greek Applied Linguistics Association* (GALA), School of Philosophy, Aristotle University of Thessaloniki, p. 209-215.
- FERGUSON Charles, 1959, “Diglossia”, *Word*, 15, p. 325–340.
- GIANDOMENICO Sica (ed.), 2006, *Open Problems in Linguistics and Lexicography*, Monza (Milano), Ed. Polimetrica
- KAGAIKE Elga and TIMUSKA Agris, 1993, “Some aspects of dialectal lexicography”, *The First International Congress of Dialectologists and Geolinguists*, Budapest, Abstracts of Scholarly Papers, p. 53-54.
- KARYOLEMOU Marilena, 1999, “Descriptive linguistics and linguistic reality: the description of dialectal communities”, *Proceedings of the 3rd International Conference on Greek Linguistics*, Athens, Ellinika Grammat, p. 596-604.
- KARYOLEMOU Marilena and PAVLOU Pavlos, 2001, “Language attitudes and assessment of salient variables in a bi-dialectal speech community”, in Fontana J., McNally L., Turell T. and Valldurí E. (eds.), *Proceedings of the First*

International Conference on Language Variation in Europe (INCLaVE1), Barcelona, Universitat Pompeu Fabra, p. 110-120.

- KATSOYANNOU Marianna, 2006, "The dictionaries of Cypriot Greek", Paper presented at the XXX SILF Annual Conference, Nicosia.
- MACENTEE-ATALIANIS Lisa and POULOUKAS Stavros, 2001, "Issues of identity and power in a Greek-Cypriot community", *Journal of Multilingual and Multicultural Development*, 22 (1), p. 19-38.
- PAPAPAVLOU Andreas, 1997, "The influence of English and its dominance in Cyprus: reality or unfounded fears", *Journal of Mediterranean Studies*, 7, p. 1-32.
- PAPAPAVLOU Andreas and PAVLOU Pavlos, 1999, "Children's language attitudes in a bidialectal setting", *Proceedings of the Third International Conference on Greek Linguistics*, Athens, Ellinika Grammata, p. 882-891.
- PAVLOU Pavlos, 1998, "A review of the sociolinguistic aspects of the Greek Cypriot Dialect", *Journal of Multilingual and Multicultural Development*, 19 (3), p. 212-200.
- PAVLOU Pavlos, 2004, "Greek Dialect use in the mass media in Cyprus", *International Journal of the Sociology of Language*, 168, p. 101-118.
- PAVLOU Pavlos, COUTSOUGERA Photini and TSIPLAKOU Stavroula (forthcoming), *The Cypriot Greek Dialect*, Munich, Lincom Europa.
- PAVLOU Pavlos and CHRISTODOULOU Niki, 2001, "Bidialectalism in Cyprus and its impact on the teaching of Greek as a foreign language", *International Journal of Applied Linguistics*, 11 (1), p. 75-91.
- SCIRIHA Lydia, 1995, *A Question of Identity: Language Use in Cyprus*, Nicosia, Intercollege Press.
- SZPILA Grzegorz, 2006, "False friends in dictionaries. Bilingual false cognates lexicography in Poland", *International Journal of Lexicography*, 19 (1), p. 73-97.
- TSIPLAKOU Stavroula, PAPAPAVLOU Andreas, PAVLOU Pavlos and KATSOYANNOU Marianna, 2006, "Levelling, koineization and their implications for bidialectalism", *Selected papers from the 3rd International Conference on Language Variation in Europe*, Amsterdam and Philadelphia, Benjamins, p. 265-276.
- WEISS Peter, 2003, *An Introduction into a (Slovenian) Dialect Dictionary*, The Slovenian Academy of Sciences.
- WEN Xiu Yang, 2007, "On pragmatic information in Learner's dictionaries, particular reference to LDOCE", *International Journal of Lexicography*, 20 (2), p. 147-173.

DEUXIÈME PARTIE

VALEURS D'EMPLOI

II.

MARQUE DIASTRATIQUE

ARG., POP., FAM. :
TROIS ŒUFS DU MÊME PANIER ?

Lola-Laurence DEVOLDER
Université Paris Descartes

Résumé — Parmi les marques dites d’usage utilisées en lexicographie afin de rendre compte de l’hétérogénéité du lexique, il en est trois – *Arg.*, *Pop.*, *Fam.* – qui ont retenu toute notre attention. Que ce soit dans une perspective stylistique ou sociolinguistique, ces marques ont fait couler beaucoup d’encre et d’aucuns supposent pouvoir les reconnaître, les définir, donc les différencier. Néanmoins, l’observation comparative des grands dictionnaires monolingues français ne laisse aucun doute : le verbe *jaspiner* sera marqué *Fam.* et *Péj.* par le *Petit Robert* 2003, *Pop.* par le *Lexis* 2002 et *Arg.* par le *Petit Larousse* 2003. Porosité des concepts ou orientations éditoriales divergentes ? C’est cet imbroglio fait de « bas langage », de « langue verte », de lexique « non conventionnel » ou simplement de « français commun », que nous tenterons de démêler. Il s’agira d’abord d’examiner les marques lexicographiques *Arg.*, *Pop.* et *Fam.*, telles qu’elles sont présentées dans les préfaces des dictionnaires et de circonscrire chacune d’elle grâce aux outils théoriques à disposition (notamment la distinction diaphasique / diastratique). Puis nous analyserons les critères (étymologique, morphologique, sémantique ou sociolinguistique) qui déterminent l’attribution de chacune de ces marques dans les dictionnaires. Enfin nous nous demanderons, au regard des réalités linguistiques actuelles, quelle pertinence trouve aujourd’hui une telle catégorisation.

Mots clés — variation, registre, lexicographie, lexicologie.

Abstract — In this paper we will attempt to clarify the terms, “low language”, “argot”, “unconventional” and simply “common French”. We will first consider the lexicographic labels, *Arg.*, *Pop.* and *Fam.*, as presented in dictionary prefaces and define each label with the theoretical tools available, in particular the sociolectal / stylistic variation distinction. Then we will analyze the criteria (etymological, morphological, semantic and sociolinguistic) that determine the allocation of each of these labels in dictionaries. Finally we ask, in light of current linguistic realities, what is the relevance of such a categorization today?

Keywords — change, register, lexicography, lexicology.

INTRODUCTION

Nous nous proposons dans cet article de revenir sur un point de la pratique lexicographique : la distribution des marques d'usage. La problématique n'est pas nouvelle. D'autres ont déjà alerté sur le difficile maniement de ces outils lexicographiques, sur les différences voire les incohérences de ce marquage (voir Messelaar 1988, Corbin 1989, Girardin 1987). Cependant, il nous semble que la discussion n'est pas close. D'abord parce que des dissonances en matière de marquage sont toujours observables entre dictionnaires et au sein d'un même dictionnaire. Également, parce que la pratique lexicographique, notamment en ce qui concerne les marques d'usage, permet un éclairage particulier de la notion de niveau ou registre de langue, ce qui nous semble une question plus linguistique qu'il n'y paraît. Enfin, parce que travailler sur les usages permet de titiller quelque peu ce que Louis-Jean Calvet (2003 : 62), avec d'autres, appelle la « linguistique de l'homogène ».

Nous avons choisi d'observer plus particulièrement la distribution des marques *Fam.*, *Pop.* et *Arg.* dans les dictionnaires monolingues français, car ce sont probablement celles qui posent le plus de difficultés dans le système lexicographique. Alors que d'aucuns les considéreront comme trois œufs du même panier, certains les traiteront comme des réalités (extra-)linguistiques bien distinctes. Dans ce contexte, notre objectif est simple : interroger la valeur opératoire de ces trois marques d'usage et, par déduction, nous intéresser à la notion de registre – ou niveau de langue –, à cette part du lexique que l'on qualifie de non standard. Autrement dit, cet article aurait pu s'appeler plus sobrement : « Les variétés du lexique non standard : de dissemblances en ressemblances ».

Précisons, bien que nous ne puissions développer ici nos prises de position terminologiques, que nous entendons par « non standard » cette part du lexique qui se voit dotée d'une marque d'usage dans les dictionnaires. Nous parlerons de lexique standard ou non marqué pour les autres entrées ; et nous qualifierons la langue constituée – ou devrions-nous dire, instituée – par les dictionnaires de langue commune. Ainsi admettons-nous le postulat selon lequel la langue commune dictionnaire se compose du lexique standard et non standard.

Notre point de départ, étant le même que celui des études citées précédemment, n'est pas très original, mais toujours aussi efficace pour la démonstration. Nous avons travaillé sur deux éditions (une des années 1970 et une des années 2000) de trois dictionnaires (*Petit Robert*, *Lexis* et *Petit Larousse*)¹. Pour éviter les questions de subjectivité dans le choix des items, nous avons relevé toutes les entrées marquées *Fam.*, *Pop.* ou *Arg.*, dans chacun de ces six dictionnaires, pour les lettres J et K². Puis, nous en avons comparé le traitement lexicographique. Ainsi, nous nous trouvons avec un échantillon de 58 entrées.

1 Dès lors PR 77, NPR 2006, Lexis 79, Lexis 2002, PL 73 et PL 2004.

2 Le choix de ces deux lettres est strictement pratique dans la mesure où il permet de présenter un inventaire exhaustif et manipulable dans le cadre d'un article. Néanmoins, les observations menées ici ont été conduites sur un corpus plus important dans le cadre de travaux davantage approfondis (Devolder 2007).

1. MARQUES D'USAGE ET HIÉRARCHISATION DU LEXIQUE

Comme nous nous y attendions, les différences de marquage d'un dictionnaire à l'autre sont importantes. Seules quatre des entrées constituant notre corpus sont marquées de façon identique par chacun des dictionnaires observés³ :

- *jeunet, te* : marqué *Fam.*,
- *jugeote* : marqué *Fam.*,
- *juteux* : marqué *Arg. milit.*,
- *kif-kif* : marqué *Fam.*

En revanche, il n'est pas rare de trouver deux ou trois marques différentes pour une même entrée :

- *jactance* : *Pop.* (PR 77 ; Lexis 79 et 2002) ; *Fam. et Vieilli* (NPR 2006) ; *Fam.* (PL 2004) ;
- *jaspiner* : *Pop.* (PR 77 ; Lexis 79 et 2002) ; *Fam. et Pél.* (NPR 2006) ; *Arg.* (PL 73 et 2004) ;
- *joint* : *Arg.* (PR 77) ; *Angli. fam.* (NPR 2006) ; *Pop.* (Lexis 79 et 2002) ; *Fam.* (PL 2004) ;
- *jouissif, ve* : *Pop.* (PR 77 ; Lexis 79 et 2002) ; non marqué (NPR 2006) ; *Fam.* (PL 2004) ;
- *junkie* : *Fam.* (PR 77 et NPR 2006, PL 73 et 2004) ; *Pop.* (Lexis 79 et 2002).

Pour l'un des items, les marques sont manifestement en totale discordance :

- *joyeuseté* : *Litt.* (PR 77 et NPR 2006) ; *Fam.* (Lexis 79 et 2002, et PL 73 et 2004).

De telles dissimilitudes de marquage s'expliquent, notamment, par l'inscription de l'entreprise lexicographique dans un projet éditorial, dont le dictionnaire se doit d'être le reflet. Chaque équipe lexicographique travaille à partir de matériaux qui lui sont propres et selon des orientations précises, conséquences de partis pris linguistiques (voire idéologiques). Aussi le système de marquage des entrées en est-il le reflet.

Cet outil métalinguistique a pour fonction, dans une perspective normalisante de reconnaissance linguistique des usages les plus divers, de légitimer l'existence de mots dans la langue en fonction de leur situation d'emploi, de leur donner force de loi et donc, à l'inverse, d'en écarter d'autres (Dubois & Dubois 1971 : 51). C'est ici toute la dimension paradoxale des marques d'usage. Certes, dans la lexicographie contemporaine, elles permettent de rompre avec une vision du lexique comme un ensemble homogène, et reflètent la diversité des productions linguistiques. Cependant, le fait même d'apposer une marque sur cette diversité induit des effets de stigmatisation.

Il est possible, par conséquent, de noter des tendances pour chacun des dictionnaires observés, répondant à des stratégies différentes, selon la tradition lexicographique dans laquelle ils s'inscrivent, le public visé, ou le positionnement plus ou moins idéologique qu'ils décident d'assumer. Nous pouvons, d'une manière générale, repérer les caractéristiques suivantes :

³ Nous ne prendrons pas en compte ici les différences de nomenclature et ne comparerons que des items présents dans au moins quatre de nos six dictionnaires.

- Le PL, quelle que soit son année de publication, est le dictionnaire qui a le plus fréquemment recours aux marques d'usage. Les entrées non marquées dans le PR et le NPR sont affublées d'une marque dans les deux PL observés, principalement la marque *Fam.* Il apparaît, en effet, que la marque *Fam.* est ici préférée à la marque *Pop.*, qui disparaît totalement de la dernière version étudiée. Quant à l'utilisation de la marque *Arg.*, elle est relativement rare.
- Les dictionnaires des éditions du Robert préfèrent utiliser la marque *Fam.* avec davantage de parcimonie, et n'ont recours que très rarement aux marques *Pop.* et *Arg.* Cette tendance s'accroît, là encore, dans les versions les plus récentes. Autrement dit, de nombreux termes marqués *Arg.*, *Pop.*, ou *Fam.* dans les autres dictionnaires se voient non marqués dans le PR et le NPR. De plus, ce sont les dictionnaires qui ouvrent le plus leur nomenclature aux items non standard, présentant alors de nombreux termes, marqués, qui sont absents des autres nomenclatures.
- Quant aux dictionnaires *Lexis*, ils utilisent les marques abondamment, à l'instar des PL. Cependant, la proportion de la marque *Pop.* est ici très importante (équivalent en nombre à la marque *Fam.*). Par ailleurs, nous constatons que le marquage, entre les deux versions observées, n'a que très peu évolué, contrairement aux autres séries de dictionnaires.

Laissons pour l'heure de côté les considérations d'ordre diachronique, auxquelles nous ferons appel plus tard dans la démonstration. Retenons simplement que ces données sont le miroir des orientations éditoriales prises par chacune des séries de dictionnaires étudiées, notamment dans leur rapport à la norme à décrire, et donc à transmettre. Les différences de marquage doivent être comprises comme une conséquence de l'aspect normatif du dictionnaire, les divers usages étant évalués en référence à des normes prises pour étalon par la position éditoriale du dictionnaire.

Cependant, la hiérarchisation du lexique n'est pas fondée sur le seul trait discriminant marqué vs non marqué. Au sein même du lexique non standard, les diverses marques, *Fam.*, *Pop.* et *Arg.*, pour celles qui nous occupent, se doivent également de refléter une organisation du lexique non conventionnel. C'est là le nœud de notre questionnement. Interroger la persistance des différences de marquage revient à questionner ce que représentent précisément ces marques lexicographiques. Comment peut-on déterminer qu'un lexème sera marqué *Fam.*, *Pop.*, ou *Arg.* ? Quels sont les critères qui permettent de considérer qu'un terme appartient au vocabulaire argotique, familier ou populaire ? Ces marques recouvrent-elles des concepts linguistiques opératoires ?

2. À LA RECHERCHE DE TRAITS DISTINCTIFS

Nous avons fait le choix d'observer chacun des items de notre corpus sous divers aspects, morphologiques, sémantiques, et du point de vue du traitement lexicographique qui en est fait. Notre objectif est de mettre en évidence des traits objectivables qui permettraient de discriminer les marques *Arg.*, *Pop.* et *Fam.* entre elles. En d'autres termes, il s'agit de rechercher s'il existe ou non des régularités linguistiques distinctives, corrélatives aux tendances de marquage⁴.

⁴ Bien que nous ayons souligné précédemment que rares sont les unités lexicales qui font l'unanimité dans leur marquage, des tendances peuvent tout de même être dégagées.

2.1 LES CARACTÉRISTIQUES FORMELLES

Le premier aspect que nous avons observé concerne les processus de création lexicale des items de notre corpus. La littérature est pléthorique concernant les procédés formels de l'argot et de la « langue » populaire, et de nombreux auteurs ont pu en établir des taxinomies. Nous retiendrons principalement Louis-Jean Calvet (1995), Pierre Guiraud (1965, 1969) et surtout Denise François-Geiger (1968) pour l'argot et Françoise Gadet (1992, 2003) pour le français populaire. Ainsi, nous disposons d'un certain nombre de critères spécifiques, nous permettant d'identifier des formes propres à l'argot d'une part, et à la langue populaire de l'autre, leurs spécificités ne se recouvrant pas tout à fait, si l'on en croit la littérature.

Les études lexicales sur le vocabulaire familier se font plus discrètes. Il n'y a guère que Gérard Petit (1998) et Jackie Schön (2002) qui se soient interrogés exclusivement sur cette part du lexique, dans une perspective davantage sémio-logique. À notre connaissance, à l'exception de quelques suffixes spécifiques⁵, aucune caractéristique formelle n'a été mise en évidence jusqu'ici.

L'observation de notre corpus montre qu'un certain nombre de caractéristiques s'avèrent formellement identifiables :

- de nombreux emprunts :
jazzy, job, joint, junkie, kid, kif ;
- des suffixations argotiques ou populaires :
jobard, jeunet, jeunot, jouailler, jouasse, jugeote ;
- des compositions souvent basées sur l'oralité :
je-m'en-fichisme, je-m'en-fichiste, je-m'en-foutisme, je-m'en-foutiste, jusqu'au-boutisme, jusqu'au-boutiste, jean-foutre ;
- des formations par reduplication :
joujou, jojo, kiki, kif-kif ;
- des processus de codage (en l'occurrence le verlan) :
keum, keuf ;
- des procédés de création sémantique par métaphore ou métonymie :
jargonner, jaspiner, jaunet, jérémiade, jeté, joyeux, jules, juteux, jeté, juter, khâgne, kiffer, kroumir ;
- des créations par dérivation :
juilletiste, jacasserie, jactance, jobarder, jobardise, jobiste, jouissif, jouissant, joyeuseté, jubilant ;
- des abréviations ou troncations :
jar (jargon), kil, kiné.

Cependant, force est de constater qu'il n'y a pas de correspondance entre ces procédés de création lexicale et l'attribution d'une marque plutôt qu'une autre. Autrement dit, ces caractéristiques de forme ne sont pas corrélables avec l'appartenance d'un terme à l'une ou l'autre des zones du lexique familier, populaire ou argotique. Partant du postulat, à l'instar des auteurs pris comme

⁵ Voir notamment Michel Roché (1999).

référence, que ces procédés sont plus productifs en lexique non standard, on peut alors envisager que nous sommes en présence de propriétés propres aux lexèmes marqués, bien plus que de traits permettant de distinguer un terme populaire d'un terme familier. Pour le moins, pouvons-nous avancer que le dictionnaire n'utilise pas les caractéristiques formelles spécifiques mises en évidence par les auteurs cités, comme critères distinctifs de nos trois marques d'usage.

2.2 LES PROPRIÉTÉS SÉMANTIQUES

Étant donnée l'hétérogénéité grammaticale de notre corpus, il nous est impossible d'établir des champs lexicaux. Aussi avons-nous simplement mis en évidence des propriétés lexico-sémantiques, qui nous permettent de regrouper nos items selon :

- les termes désignant l'individu :
jojo, jean-foutre, jeunet, jobard, jeunot, jobiste, je-m'en-fichisme, je-m'en-foutisme, jules, junkie, jusqu'au-boutiste, kroumir, keuf, joyeux ;
- des objets ou des situations du quotidien :
jaunet, job, joint, kil ;
- des activités sociales :
jargonner, jacasser, jouer, jouailler ;
- des termes relatifs aux émotions :
jubiler, kiffer, jouissant, jouissif, jérémiades, jubilation.

Au-delà de ces « regroupements sémantiques », l'observation des gloses données par le dictionnaire nous donne l'occasion d'insister sur une des caractéristiques, sans doute la plus pertinente, spécifique au lexique non standard : la composante dépréciative, péjorative qui pèse sur ces termes. En effet, la majorité des termes étudiés porte, par connotation ou dénotation, un jugement de l'ordre de la réprobation sociale, que ce jugement soit inclus dans la définition comme une appréciation axiologique (*jouailler* « jouer petit jeu », *jubilant* « joie vive et excessive », *jobard* « très naïf, qui se laisse duper facilement »), ou qu'il soit lié directement au référent (*kil* « litre de rouge », *jouissif* « qui fait jouir », *joint* « cigarette de marijuana », *jules* « homme du milieu », *junkie* « toxicomane »).

Cependant, là encore, les indices lexico-sémantiques ne peuvent pas constituer des critères distinctifs des vocabulaires argotique, populaire, ou familier, mais semblent au contraire fonctionner comme facteur homogénéisant du lexique non standard.

2.3 LE RAPPORT À LA NORME

Enfin, nous avons étudié les procédures définitoires utilisées par le dictionnaire, afin de voir si elles constituent, ou non, un critère distinctif des marques attribuées. Cette démarche peut apparaître surprenante, néanmoins il nous semble que la définition lexicographique informe sur le rapport qu'établit le discours métalinguistique du dictionnaire, entre lexique marqué et non marqué. En cela, nous avons considéré les définitions comme des objets de discours présentant une équivalence de signes entre vocabulaire standard et non standard. Il s'agit donc d'observer comment le dictionnaire décrit le vocabulaire *Fam.*, *Pop.* et *Arg.*, et de voir si ce rapport à la norme peut être, ou non, considéré comme indice de différenciation dans la pratique lexicographique.

Deux procédures définitoires sont principalement utilisées :

- La définition par quasi-synonyme : le dictionnaire monolingue fonctionne alors comme un dictionnaire bilingue et offre, en quelque sorte, une traduction d'une « langue » non conventionnelle par une « langue » normée. Le terme marqué s'avère alors un « autre du mot » standard pour reprendre la terminologie de Petit (1998), une sorte de doublon. Le lexique normé assure un rôle d'interprétant du lexique marqué, dans la mesure où l'on n'accède au référent de l'entrée marquée que par le référent du terme non marqué qui le définit :

jactance « bavardage »,

jar « argot »,

jeté « fou, cinglé »,

keuf « policier ».

- La définition logique ou par inclusion : le dictionnaire utilise pour définir le mot marqué une base (quasi-)hypéronymique doublée d'une spécificité. L'unité marquée est ici considérée comme un « autre mot », un autre concept, par extension :

jojo « enfant turbulent »,

jean-foutre « personne incapable, sur qui on ne peut pas compter »,

jacasserie « bavardage continu et bruyant »,

jérémiade « plainte sans fin, qui importune ».

Sans surprise, ces deux types de procédés définitoires se répartissent équitablement sur l'ensemble de notre corpus et ne sont pas préférés pour une marque plutôt qu'une autre. À nouveau, les procédures définitoires ne peuvent être prises en compte comme critère déterminant l'attribution de telle ou telle autre marque. Ainsi, pouvons-nous infirmer l'hypothèse, largement répandue, selon laquelle l'argot serait un vocabulaire qui « double » le français standard, alors que le populaire en serait la face expressive, imagée, spécifiée.

Ces quelques points d'observation d'un corpus dictionnaire nous ont permis de mettre en évidence (ou simplement de confirmer) des caractéristiques propres au lexique non standard, tant au niveau du signifiant que du signifié. Contrairement à ce que l'on peut constater notamment dans la préface du PR 77 (p. xvij), ce n'est pas le mot, la forme linguistique – en tout cas pas seulement – qui est qualifiée par la marque, mais aussi la notion, la chose, le concept.

En revanche, nous n'avons pu repérer les traits pertinents qui permettraient de distinguer les marques *Pop.*, *Fam.*, et *Arg.* Les paramètres que nous avons étudiés se révèlent insuffisants pour déterminer l'attribution d'une marque d'usage plutôt qu'une autre. La délimitation de ces zones lexicales est bien moins aisée que ne semble l'établir le système de marque lexicographique.

Revenons alors à la source, et observons comment les lexicographes eux-mêmes, dans leurs préfaces, justifient l'existence de ces trois marques.

3. ARG., POP., FAM. ET LE MÉTADISOURS LEXICOGRAPHIQUE

Malheureusement – mais si tel n'était pas le cas, le questionnement perdrait son intérêt –, les marques *Fam.*, *Pop.*, et *Arg.* ne sont pas les marques d'usages les plus détaillées, bien au contraire.

Les définitions ou descriptions de ces outils ne sont pas mentionnées dans les préfaces des PL. Quant aux *Lexis*, ils expliquent que seront présentés les termes du vocabulaire courant, en les pourvoyant éventuellement d'une indication de *niveau de langue* précisant la situation de communication où ils sont employés. Une parenthèse signale que par *niveau de langue* nous devons comprendre langue familière, populaire, ou au contraire langue soutenue, littéraire. Aucun élément ne permet donc de différencier ces marques.

C'est finalement Alain Rey qui, pour le PR et le NPR, apporte les précisions les plus éclairantes. Tout d'abord il stipule, dans la préface du PR 77, que les marques d'usage caractérisent la valeur de l'emploi, dans le temps (Vx.), dans l'espace (Région.), par la fréquence (Rare), dans la société (*Fam.*, *Pop.*). Il précise alors qu'il faut bien distinguer, d'une part, les usages qui constituent de véritables signaux d'appartenance sociale comme *Arg.*, ou *Pop.* et, d'autre part, *Fam.*, qui indique la situation de communication.

Cette distinction sera développée dans le tableau des abréviations (p. xxv) qui fait suite à la préface, et qui propose de véritables gloses pour chacune des marques étudiées :

Arg. : mot d'argot, emploi argotique limité à un milieu particulier, surtout professionnel, mais inconnu du grand public ; *Arg. fam.* : mot d'argot passé dans le langage familial ; argotique.

Fam. : familial. Usage parlé et même écrit de la langue quotidienne : conversation, etc., mais ne s'emploierait pas dans les circonstances solennelles ; concerne la situation de discours, non l'appartenance sociale, à la différence de *Pop.*

Pop. : populaire. Qualifie un mot ou un sens courant dans la langue parlée des milieux populaires (souvent argot ancien répandu), qui ne s'emploierait pas dans un milieu social élevé (à distinguer de *Fam.* qui, lui, concerne la situation de communication).

Ces indications lexicographiques nous renseignent alors sur deux points essentiels. D'une part les marques *Pop.* et *Arg.* fonctionnent comme des marqueurs sociaux (appartenance de classe ou de groupe), contrairement à la marque *Fam.* qui est une indication situationnelle (réduction de la distance interlocutive). D'autre part ces aires de vocabulaire ne sont pas étanches, certains termes argotiques pouvant être considérés comme *Fam.* et d'autres comme *Pop.* : certains mots d'argot étant « passés dans le langage familial », et certains termes populaires étant de l'« argot ancien répandu ».

C'est précisément à cette perméabilité que se heurtent les lexicographes dans l'étiquetage des entrées et qui rend la tâche si difficile. Une solution consisterait à évaluer, par enquête, si une entrée dictionnaire est préférentiellement considérée ou employée comme un marqueur social ou un marqueur situationnel. Et si elle est un *signum* social, est-elle celui d'un groupe ou bien celui d'une classe sociale, en l'occurrence, la classe dite populaire. Nous percevons immédiatement l'impossibilité d'une telle réalisation qui serait, sitôt achevée, devenue obsolète. Le système dictionnaire ne peut échapper à la subjectivité, donc à l'aléatoire, à l'intuition, dans le domaine du marquage des mots. Il revient, au lexicographe, cette impossible tâche qui consiste à catégoriser des usages, des segments de discours, c'est-à-dire à figer ce qui est, par définition, dynamique.

Nous continuons alors notre exploration du côté des études en description des langues, et plus précisément en sociolinguistique, sans doute plus à même que le dictionnaire de rendre compte de la réalité des usages linguistiques.

4. LES ÉVALUATIONS DES PRATIQUES LANGAGIÈRES

Les enquêtes, réalisées sur l'argot et la langue populaire, ne semblent pas confirmer les hypothèses proposées par les préfaces des dictionnaires, qui présentent l'argot comme un vocabulaire crypté, le populaire comme un marqueur de classe, et le familier comme un indicateur situationnel.

Déjà, Denise François (1989) affirmait que l'argot avait perdu son caractère crypto-ludique ; sa principale fonction était de créer une connivence et de « rendre le discours plus familier ». Elle développe cette interprétation de l'argot comme *signum* social de second degré, dépassant les stratifications sociales : « on s'affirme affranchi en matière de langage toutes classes sociales plus ou moins confondues » (François 1989 : 94). Françoise Gadet reprend également cette thèse dans ses articles les plus récents à propos de la langue des jeunes. Elle parle de « socialisation en réseaux serrés » (Gadet 2003 : 111). Elle considère alors la notion de français populaire comme plus interprétative que descriptive, les critères d'argotité ou de popularité étant difficiles à établir de façon rigoureuse et systématique.

Cette observation des usages sociaux est confirmée par les récentes enquêtes en lexicologie sociale (Lesigne 2000) effectuées auprès de jeunes lycéens, qui montrent qu'ils emploient de moins en moins de termes marqués socialement ou géographiquement. C'est également le sens des travaux d'Anthony Lodge (1992) réalisés sur les habitudes langagières de jeunes issus de quartiers défavorisés et dont les conclusions remettent clairement en cause la corrélation entre vocabulaire populaire et appartenance de classe.

Il semble en effet qu'il n'y ait plus vraiment aujourd'hui d'usages constatés de termes employés par certaines classes de la société (notamment les plus populaires) et inconnues des autres locuteurs : pour preuve les termes considérés principalement comme *Pop.* dans notre corpus : *jacter*, *joint*, *junkie*, etc. Ils peuvent, pour certains, être rarement employés, mais ils ne sont pas ignorés des locuteurs instruits. La fonction discursive de l'usage de termes argotiques ou populaires est aujourd'hui décrite par les linguistes et sentie, par les usagers, comme des usages familiers, c'est-à-dire qu'elle s'apparente à une réduction de la distance interlocutive, qu'elle vise à créer une certaine familiarité entre locuteurs.

C'est également ce que l'on peut remarquer quand on observe la dimension diachronique que présente notre corpus. Pour chacun des trois dictionnaires, même si la tendance est moins marquée pour le *Lexis* (dont le marquage des termes n'évolue pas beaucoup sur les trente dernières années), il est très nettement repérable qu'un nombre important d'entrées marquées *Pop.* ou *Arg.*, dans les éditions les plus anciennes, sont marquées *Fam.* dans les plus récentes.

jacter : *Pop.* (PR 77), *Fam.* (NPR 2006),

je-m'en-fichisme : *Pop.* (PL 73), *Fam.* (PL 2004).

Quant aux unités fraîchement intégrées à la nomenclature, elles sont directement marquées *Fam.*, y compris si leur forme est nettement identifiable : pour exemple, *keuf* et *keum*, qui sont du verlan, donc formellement associables à

l'argot, sont marqués *Fam.* Cette tendance qui vise à privilégier la marque *Fam.* est en fait liée à l'évolution des pratiques sociales, et donc linguistiques, qui tendent à se niveler, notamment par le biais de la mobilité sociale, l'éducation de masse et ce que l'on appelle les mass-média. Cela passerait un peu comme une lapalissade, tant il est vrai que dans une société où l'on cherche de plus en plus à neutraliser les différences de classes sociales, il est effectivement assez évident que ne perdurent pas des usages de classe ou des usages cryptiques.

Nous avançons alors que les particularités linguistiques non standard relèvent aujourd'hui bien plus des modalités d'interaction que du social⁶. Le marquage lexicographique, plus justement que ne le sont les préfaces des dictionnaires, fait écho aux récents écrits sociolinguistiques sur l'argot et la langue populaire, et tend à ne plus considérer ces catégorisations du lexique comme des concepts opérationnels. Toute la difficulté du lexicographe, que nous avons repérée comme étant liée à la distinction entre diaphasique et diastratique, n'est finalement pas opératoire pour guider le marquage des entrées, dans la mesure où elle est bien plus théorique que pratique.

Ainsi, au terme de nos observations, pouvons-nous avancer que, d'un point de vue linguistique et sociolinguistique, les distinctions traditionnelles entre formes « argotiques », « populaires » et « familières » n'ont plus de réalité fonctionnelle. En revanche, la césure opposant usages normés et hors la norme, standard et non standard, persiste et garde toute sa pertinence.

5. PERSPECTIVES LEXICOGRAPHIQUES

D'un point de vue lexicographique, nous avons montré en quoi la multiplicité des marques *Fam.*, *Pop.* et *Arg.* n'est pas fondée au plan strictement linguistique. Elle l'est encore moins pour l'utilisateur du dictionnaire, *a priori* non averti, et tend plutôt à en brouiller la compréhension. Nous soutenons que la seule distinction réellement fonctionnelle est celle qui oppose les entrées marquées aux non marquées, le lexique standard au non standard.

Par conséquent, il nous semble qu'une seule marque pourrait convenir au marquage lexicographique des entrées non standard. Nous avons conscience de l'audace de cette proposition, mais il nous semble qu'il s'agit simplement d'entériner une situation existante et constatée. Déjà au début des années 1970 Josette Rey-Debove (1971 : 93) écrivait :

Dans l'état actuel du langage, on peut considérer qu'il y a une dichotomie essentielle, qui prime sur les autres : le langage familial employé par tous dans les relations ordinaires entre égaux (langage écrit, parlé, littéraire) et le langage officiel employé dans les relations officielles ou hiérarchiques ; chacun de ces langages comporte plus ou moins de termes marqués selon les individus.

La dichotomie marqué / non marqué était déjà considérée comme principale. Nous ajoutons simplement, pour renforcer l'argumentation, qu'elle est d'autant plus prégnante qu'en discours les emplois relevant des dimensions diaphasique et diastratique produisent les mêmes effets.

En conséquence, les marques *Pop.* et *Arg.* seraient réservées à des usages bien spécifiques, et alors considérées comme des marques diatechniques (usages

⁶ Précisons que nous nous situons au niveau du lexique, et qu'il n'en est peut-être pas de même pour ce qui est de la syntaxe.

restreints à des groupes identifiables, relativement cryptiques, qui auraient leur place dans un dictionnaire généraliste notamment pour comprendre la littérature spécialisée), ou bien elles deviendraient des indicateurs étymologiques sur le modèle *Arg. Fam.*⁷, ou plus clairement « D'abord *Arg.* puis *Fam.* » ou « D'abord *Pop.* puis *Fam.* » Enfin, corrélativement, il faudrait ne pas oublier de préférer l'appellation *marque d'usages* au pluriel, plutôt que l'actuelle *marque d'usage* au singulier.

RÉFÉRENCES BIBLIOGRAPHIQUES

- COHEN Marcel, 1970, « *C'est rigolo* n'est pas populaire », *Le Français Moderne*, 1, p. 1-9.
- CALVET Louis-Jean, 1995, *L'Argot*, Paris, Puf.
- CALVET Louis-Jean, 2003, « L'argot et la "langue des linguistes" : des origines de l'argotologie aux silences de la linguistique », *Marges linguistiques*, 6, p. 55-64.
- CORBIN Pierre, 1989, « Les marques stylistiques / diastratiques dans le dictionnaire monolingue », dans F. J. Hausmann *et alii*, *Wörterbücher, ein internationales Handbuch zur Lexikographie I*, Berlin u. New-York, Walter De Gruyter, p. 673-680.
- CORBIN Pierre, 1991, « Le maquis lexicographique. Aperçus sur l'activité lexicographique monolingue dans le domaine français à la fin du XX^e siècle », *Le français moderne*, 94, p. 6-26.
- DEVOLDER Laurence, 2007, *Langue et registre(s) : illustration à partir de l'indice d'usage familial*, thèse de l'Université Paris Descartes (Paris V).
- DUBOIS Jean et Claude, 1971, *Introduction à la lexicographie*, Paris, Larousse.
- FRANÇOIS-GEIGER Denise, 1968, « Les argots », dans A. Martinet (éd.), *Le Langage*, Paris, Gallimard, p. 620-646.
- FRANÇOIS-GEIGER Denise, 1989, « Les paradoxes des argots », *L'Argoterie*, Paris, Centre d'Argotologie de l'UER de linguistique de l'Université Paris V, p. 91-104.
- GADET Françoise, 1992, *Le Français populaire*, Paris, Puf.
- GADET Françoise, 2003, *La Variation sociale en français*, Paris, Ophrys.
- GIRARDIN Chantal, 1987, « Système de marques et connotations sociales dans quelques dictionnaires culturels français », *Lexicographica*, 3, p. 76-102.
- GLATIGNY Michel, 1998, *Les Marques d'usages dans les dictionnaires français monolingues du XIX^e siècle*, *Lexicographica*, Series maior 91, Tübingen, Max Niemeyer Verlag.
- GUIRAUD Pierre, 1965, *Le Français populaire*, Paris, Puf.
- GUIRAUD Pierre, 1969, *L'Argot*, Paris, Puf.
- LESIGNE Hubert, 2000, *Les Banlieues, les profs, les mots*, Paris, L'Harmattan.
- LODGE Robert Anthony, 1992, « Le vocabulaire non standard suivant les perceptions des locuteurs actuels », dans GEHLF (éd.), *Grammaire des fautes et français non conventionnel, Actes du IV^e Colloque international de l'École Normale Supérieure*, Paris, Presse de l'ENS, p. 341-353.

⁷ Proposition déjà établie par Marcel Cohen (1970 : 1).

- MESSELAAR Petrus Adrianus, 1988, « Les marques *familier* et *populaire* envisagées d'un point de vue lexicologique et lexicographique », *Cahiers de lexicologie*, 53, p. 91-106.
- PETIT Gérard, 1998, « Approche lexicale et sémantique du vocabulaire familier », *Cahiers de lexicologie*, 72, p. 5-40.
- REY-DEBOVE Josette, 1971, *Étude linguistique et sémiotique des dictionnaires français contemporains*, La Haye, Mouton.
- ROCHÉ Michel, 1999, « Les suffixes évaluatifs comme marqueurs d'approximation », *Sillexicales*, 2, p. 209-217.
- SCHÖN Jackie, 2002, « Pour un traitement systématique des acceptions familières dans les dictionnaires », *Romania sans frontière*, 2, p. 605-615.

L'EXPRESSIVITÉ ET LA MARQUE LEXICOGRAPHIQUE :
ÉTUDE COMPARATIVE FRANCO-TCHÈQUE
D'UN CORPUS DU LEXIQUE NON STANDARD.

LES MARQUES *FAM.*, *POP.*, *ARG.* VS *EXPRESSIVITÉ* EN
LEXICOGRAPHIES FRANÇAISE ET TCHÈQUE

Alena PODHORNÁ-POLICKÁ
Université Masaryk de Brno (République tchèque),
Laboratoire Dynalang (Paris V)

Résumé — La confusion terminologique relative à la définition de l'argot dans les linguistiques française et tchèque est le reflet d'une confusion plus profonde concernant le classement du lexique non standard. De plus, la dynamique de l'usage de ce lexique fait que cette question échappe incessamment à une analyse complexe en temps réel, problème qui s'ajoute aux divergences de classement entre lexicographes et lexicologues. Il suffit de regarder deux dictionnaires différents ou même deux éditions d'un même dictionnaire pour constater que les limites entre les marques *vulg.*, *arg.*, *pop.* et *fam.* ne sont pas du tout étanches en lexicographie française. La lexicographie tchèque n'est pas non plus épargnée par ce genre de critique, mais grâce à l'insertion de la marque *expr.*, « expressive », elle arrive à contourner un bon nombre de problèmes liés à la catégorisation des cas-limites. En étudiant un corpus d'une cinquantaine de mots d'origine argotique, le changement de leurs marques lexicographiques dans le *Petit Larousse* et le *Petit Robert* entre les années 1960 et 1990 et leurs équivalents dans les dictionnaires tchèques, nous allons essayer de commenter les différentes approches du classement de ce type de lexique dans les deux systèmes lexicographiques observés.

Mots clés — français sub-standard, langue tchèque, marque *arg. fam.*, marque d'expressivité.

Abstract — The terminological confusion over the definition of slang in French and Czech linguistics reflects a deeper confusion in the classification of the non-standard lexicon. Moreover, due to the dynamics of lexical usage this matter constantly eludes complex analyses in real time. This problem is further compounded by the different classification systems of lexicographers and lexicologists. One has simply to look at two dictionaries – or even two different editions of the same dictionary – to see that the boundaries between the labels

vulg., *arg.*, *pop.* and *fam.* are not at all watertight in French lexicography. Czech lexicography is also open to this kind of criticism, but its insertion of the expr. “Expressive” allows it to circumvent many problems associated with the categorization of borderline cases. By studying a corpus of fifty words classified as argot / slang, the changes in their lexicographical labeling in *Le Petit Larousse* and *Le Petit Robert* between 1960 and 1990, and their equivalents in Czech dictionaries, we will comment on different approaches to the classification of this type of word in the two lexicographical systems examined here.

Keywords — sub-standard French, Czech language, *arg.* and *fam.* labels, expressiveness.

INTRODUCTION

La confusion terminologique qui s’attache à la définition de l’argot, au sens moderne du terme, non seulement en linguistique française, mais également en linguistique tchèque, est le reflet d’une confusion plus profonde concernant le classement du lexique non standard. Les critères non univoques des lexicographes et les reformulations des classements en matière de lexicologie selon des points de vue variables font que cette question échappe incessamment à une analyse complexe en temps réel, sans oublier la dynamique de l’usage des expressions du non standard. La linguistique tchèque traite ce lexique (*nespisovná slovní zásoba*) dans une optique relativement divergente de celle de la linguistique française, même si une notion-clé est commune à ces deux linguistiques : il s’agit du terme de *lexique marqué* – équivalent de « *příznakové lexikum* » – en opposition au *lexique neutre* – « *nepříznakové / neutrální lexikum* ». Notons qu’en France, la catégorisation bipartite des lexèmes en *neutres* et *marqués* permet de contourner les problèmes liés à la classification ambiguë entre les *niveaux* et les *registres*.

En étudiant un corpus d’une cinquantaine de mots d’origine argotique, le changement de leurs marques lexicographiques dans le *Petit Larousse* (PL) et le *Petit Robert* (PR) entre les années 1960 et 1990 et leurs équivalents dans les dictionnaires tchèques, nous allons essayer de commenter les différentes approches du classement de ce type de lexique dans les deux systèmes lexicographiques observés¹.

Rappelons brièvement les traits communs dans les deux systèmes du marquage lexicographique. Les lexèmes marqués, en tant qu’entrées dans les dictionnaires, comportent une *marque métalinguistique* ou bien une *marque d’usage / d’emploi* – équivalent de « *stylistický / stylový příznak* » (« *marque stylistique / de style* ») en linguistique tchèque, c’est-à-dire un indice lexicographique qui marque, à côté du sens dénotatif, des valeurs sémantiques de nature connotative qui servent à préciser les restrictions d’emploi du lexème².

1 Le corpus épilinguistique a été composé en France dans des lycées professionnels à Paris et à Yzeure (Allier), en République tchèque, dans un lycée professionnel à Brno. L’approche sociolinguistique a reposé sur l’observation participante de groupes de jeunes concernant leurs pratiques langagières et sur des entretiens semi-directifs.

2 Pour ce qui concerne les chevauchements de sens autour de la notion de « *stylistique* » en linguistique tchèque et française, voir Podhorná-Polická (2007 : 29-35 et 183-190).

Ces restrictions peuvent être divisées en deux types : les unes, relativement objectives et communes pour les deux linguistiques sont des marques qui indiquent les restrictions spatiales concernant les régionalismes, les restrictions temporelles (un mot peut être marqué comme archaïque / vieux, vieilli ou moderne), les restrictions d'emploi dans un domaine de savoir particulier touchant aux mots scientifiques et aux professionnalismes (technolectes), etc. Le deuxième type de restriction pose problème à cause de la subjectivité d'attribution de la marque et de sa variabilité selon les approches adoptées : il s'agit des marques indiquant la provenance sociale (argotique, populaire³) et des marques stylistiques / situationnelles (littéraire, soutenu ou familier⁴, vulgaire, péjoratif / injurieux, etc.). Ce sont ces marques, dénommées « marques s/d » (*stylistiques / diastratiques*) par Corbin (1989), qui feront l'objet de notre analyse. En linguistique tchèque, les quarante ans du communisme, pendant lesquels l'existence de couches sociales a été officiellement niée, ont provoqué la domination de l'approche diaphasique dans le traitement du lexique sub-standard. De ce fait, nous constatons une distorsion moins importante entre le traitement lexicographique tchèque des marques d'ordre diaphasique et d'ordre diastratique⁵.

1. FONCTIONNALITÉ DES MARQUES LEXICOGRAPHIQUES SUB-STANDARD FRANÇAIS : TOUR D'HORIZON

La confusion entre les marques métalinguistiques résulte du fait que la qualification des mots du sub-standard s'opère selon les jugements subjectifs que les lexicographes portent sur des usages, mais également sur la provenance des mots connotés socialement. Tandis que la marque « familier » fait référence avant tout à une variété discursive de français parlé dans un certain type de situations communicationnelles, s'inscrivant dans le cadre familial ou dans celui des relations amicales, les marques « populaire » et « argotique » renvoient tout simplement à l'origine sociale des mots. Les marques « populaire » et « argotique » nous semblent alors être de vrais fossiles lexicographiques si l'on adopte une approche purement fonctionnelle. Seule la marque « familier » est débarrassée de sa connotation sociale, c'est pourquoi elle paraît avoir le plus d'avenir. Dans cet article, notre objectif est de passer en revue toutes ces « marques s/d » sub-standard et de les mettre en contexte diachronique et comparatif vis-à-vis de la pratique lexicographique tchèque.

Le traitement des « marques s/d » par les dictionnaires se heurte à la pluralité des objectifs : « s'y entremêlent en effet des visées sociolinguistiques (pointer

3 L'équivalent de la marque « populaire » n'existe pas en lexicographie tchèque, car la variation diastratique n'est pas prise en compte sauf pour le cas de l'« argot » (à distinguer du « slang »).

4 L'équivalent de la marque « familier » semble être, en linguistique tchèque, la marque « hovorové », (équivalent du « colloquial » anglais), parfois aussi « obecně české » (relevant du tchèque commun), ou bien même « expresivní » (expressif). Il est regrettable que le terme « colloquial » soit absent de la nomenclature française ; il nous semble que l'avantage de ce terme est qu'il renvoie à la *conversation* et non à la *famille* : les apprenants de FLE font régulièrement la confusion entre la marque *fam.* « familier » française et *fam.* « familiární » tchèque, cette dernière étant attribuée à des termes affectifs concernant le foyer familial (du type *faire dodo*, *faire pipi* ou divers hypocoristiques).

5 Même si, à l'époque post-communiste, on a assisté à des discussions linguistiques sur la raison d'être de la seule marque purement diastratique, à savoir *arg.* (au sens du lexique à visée cryptique utilisé par des sous-groupes défavorisés, voire délinquants), voir à ce propos Jaklová (1999).

des usages socialement marqués), stylistiques (repérer l'inscription lexicale des conditions d'énonciation) et de savoir-vivre (énoncer des jugements de bienséance) » (Corbin 1989 : 673). En somme, les défauts du système de marquage français sont critiqués par les linguistes français depuis un certain temps et ont d'ailleurs fait l'objet du colloque qui est à l'origine de cette publication.

Adoptons donc une approche pragmatique, celle d'un usager du *Robert* ou du *Larousse* (et notamment un usager étranger). Pour lui, le plus important est de trouver des indications sur la pragmatique du sens. Ainsi, il s'agit de mettre en évidence un indice permettant de décider si le mot, dans son usage le plus fréquent, est considéré comme neutre / standard, s'il est limité à un emploi familier ou bien péjoratif (la marque « vulgaire » est à traiter à part) et surtout, s'il a, dans son emploi le plus fréquent, une valeur affective, emphatique (un usager des dictionnaires tchèques recevra cette information grâce à la marque *expr.*).

La stratification sociale dans la société française actuelle n'implique plus l'imperméabilité quant au choix lexical. Le recours à des mots familiers peut agir comme un signe de connivence. Claire Blanche-Benveniste (2000 : 54) note à ce propos que les mots appartenant au « français non conventionnel » tels que *bouquin, flic, gaffe, pieu, se pieuter, roupiller, bouffer, bâfrer, picoler, rigoler, se marrer, se planter* ne sont plus limités à l'argot⁶. Il semble évident qu'à notre époque, les indications sur la provenance sociale du mot semblent jouer un rôle plutôt marginal. C'est sans doute la raison pour laquelle les dictionnaires d'usage tels que le PR ou, dans une moindre mesure, le PL ont quasiment abandonné la marque « populaire » en faveur de la marque « familier », ce que nous allons montrer plus loin dans cet article⁷. Or, la marque *fam.* a également des connotations restrictives puisqu'elle ne dissocie pas suffisamment les termes familiers hypocoristiques et les termes d'origine argotique, passés à l'argot commun. Le PR, par exemple, gère cette bifurcation sémantique par l'introduction d'une nouvelle marque qui rassemble deux marques déjà connues, *arg.* et *fam.* En somme, il apparaît que les marques d'emploi dans les dictionnaires français, notamment celles qui renvoient à des niveaux / registres de langue sub-standard (qui eux seuls sont au cœur de notre recherche), traversent actuellement une crise profonde.

La comparaison avec la pratique lexicographique tchèque pourrait donc être utile pour résoudre le problème des chevauchements des marques françaises. Nous sommes d'avis qu'en redéfinissant le contenu sémantique de la marque « argotique » grâce à la notion d'*expressivité*, le problème du marquage lexicographique pourrait être simplifié.

L'expressivité lexicale est une notion bien ancrée dans la terminologie linguistique tchèque. La notion d'expressivité chez Bally (1921 [1909]) a été développée dans le cadre de la lexicologie (Machek 1930, Zima 1963) au sein du Cercle linguistique de Prague, pour aboutir à une théorie de l'expressivité lexicale. J. Zima a introduit une classification de l'expressivité, désormais communément acceptée, qui distingue :

- une expressivité inhérente (le mot est expressif même hors contexte) ;

6 Dans l'argotologie moderne, ce même lexique est désigné comme « argot commun » (voir plus loin François-Geiger 1989).

7 Voir des lexèmes qui n'ont pas été mis en italiques dans le tableau en annexe.

- une expressivité adhérente (l'expressivité est présente dans certains contextes uniquement – c'est par exemple souvent le cas de l'expressivité métaphorique) ;
- une expressivité stylistique, appelée expressivité contextuelle, qui joue sur l'actualisation du discours par le biais de la transgression des registres dans un but impressif.

Une comparaison avec la sémantique interprétative de François Rastier (1987) s'impose : ce dernier distinguait les sèmes inhérents, afférents ainsi que les sèmes actualisés et neutralisés – ce qui fait un parallèle plus général à la typologie tchèque de l'expressivité.

En effet, pour la lexicographie tchèque, la marque *expr.* s'utilise de manière tout à fait officielle, notamment pour le lexique où l'expressivité est inhérente, voire adhérente (pour des acceptions polysémiques). Dans ce qui suit, nous tenterons de circonscrire les marques lexicographiques françaises, en fonction de leur rapport à l'argot dans son acception moderne et vis-à-vis de la notion d'expressivité. Nous espérons apporter, grâce à cette approche comparative, un humble témoignage de l'importance d'une unification conceptuelle du marquage français et du rôle de l'argotologie moderne pour la consolidation définitoire des marques s/d.

2. ARGOT VS MARQUE ARGOTIQUE

Parmi les marques lexicographiques françaises, la marque « argotique » semble être la plus difficile à discuter, puisque sa définition est aussi ambiguë que la définition même de l'argot. Denise François-Geiger a proposé de distinguer l'*Argot* et les *argots*, terminologie adoptée désormais par l'argotologie moderne (François-Geiger 1991 : 5). En effet, l'*Argot*, au singulier et avec une majuscule, signifie l'argot des malfaiteurs, c'est-à-dire l'argot dans son sens classique, tandis que si l'on parle des *argots* au pluriel, c'est dans le sens moderne du terme qui prend en compte la variation lexicale dans des milieux cohésifs, dont les membres sont unifiés soit autour d'une activité commune – d'où la proximité des notions de jargon et de jargot (v. Sourdou 1991 et 2002), soit autour d'un sentiment identitaire communautaire (mots identitaires générationnels – argot des jeunes, socio-spatio-ethniques – argot des jeunes des cités, etc.). Telle est la vision des argotologues.

Or, le public non spécialisé, et même parfois les chercheurs d'autres disciplines, ont des idées imprécises et, du coup, très confuses sur le contenu définitoire de l'argot. En effet, plusieurs linguistes (Walter 1988 : 314-315, Jolin-Bertocchi 2003 : 76) expliquent qu'on peut habituellement retrouver trois types de lexiques très différents sous le terme englobant d'« argot » :

Tout d'abord, il s'agit de l'ancien argot des malfaiteurs, des voleurs, des bagnards et des prisons qui correspond globalement à la marque *arg.*, « argotique », dans la tradition lexicographique tchèque.

Ensuite, des argots « modernes » basés sur une fonction conviventielle et identitaire qui s'établit entre personnes d'un même milieu (professionnel, sportif, scolaire, militaire, carcéral, etc.)⁸ – dans les dictionnaires tchèques, ce

⁸ On peut également classer dans ce type d'argot l'argot des jeunes ou celui du français contemporain des cités (Goudaillier 2002).

type de lexique est étiqueté par la marque *slang.*, « slanguisme », éventuellement avec la précision du milieu (*sport slang.*, « argot sportif », etc.).

Enfin, il s'agit de l'« argot commun »⁹, c'est-à-dire du *vocabulaire non conventionnel* qui peut aussi bien être désigné comme *populaire* ou bien également, selon la tendance contemporaine, comme *familier*, mais aussi comme *vulgaire* si le mot transgresse les tabous – tandis que la tradition anglo-saxonne étiquetterait ce type de lexique sans hésitation sous le nom de *slang* (v. Lodge 1997 : 2) ; c'est à ce niveau justement que les dictionnaires tchèques opèrent avec la marque *expr.*, privilégiant ainsi l'aspect expressif et conniventiel, au détriment des aspects étymologiques.

3. L'EXPRESSIVITÉ : ABSENCE TERMINOLOGIQUE MAIS PRÉSENCE IMPLICITE DANS LES COMMENTAIRES LEXICOGRAPHIQUES

Afin de prouver la faible fonctionnalité des marques métalinguistiques pour les niveaux du sub-standard du français, nous proposons maintenant une brève analyse de la dynamique des marques lexicographiques dans les deux dictionnaires les plus usités – le PR et le PL – en comparant une édition des années 1960 et une édition des années 1990, soit une trentaine d'années plus tard. Nous avons procédé à l'analyse de 50 expressions qui figurent dans le *Dictionnaire de l'argot* de chez Larousse (Colin, Mével et Leclère 1992, désormais DA), et qui y sont répertoriées en tant qu'expressions d'origine argotique. La sélection a été aléatoire (à partir de notre corpus d'argotismes notés par les élèves enquêtés), mais pourtant relativement consciente dans la mesure où nous avons cherché à sélectionner un échantillon d'expressions communément connues, dont l'usage actuel est le plus varié possible.

Tous ces mots déboulent le lexique standard par leur connotation axiologique et leurs traductions en tchèque comportent le plus souvent la marque *expr.* Si ces expressions figurent dans un dictionnaire d'argot, elles devraient, en principe, être toutes suivies de la marque *arg.* Or, le DA lui-même complète certains termes par des notes qui précisent la pragmatique de leur usage actuel¹⁰. Les auteurs expliquent cette situation ainsi :

Nous avons cru bon de faire quelques remarques concernant l'emploi des mots, en particulier pour souligner la coappartenance du mot à plusieurs registres : populaire et argotique ou familier et argotique, ou pour signaler l'évolution historique de tel mot, dont la carrière a commencé dans l'argot « fort », pour passer peu à peu dans un domaine plus vaste, celui d'une familiarité courante : glissement intuitivement perceptible, mais évidemment malaisé à repérer de façon précise. (DA : xiiij).

L'argot n'est plus considéré comme le langage cryptique des malfaiteurs, hermétiquement fermé aux non-initiés. Il évolue parallèlement au lexique standard et il est donc nécessaire d'actualiser ces connotations stylistiques à l'aide

9 Notion promue par Denise François-Geiger (1989 : 84) et couramment utilisée en argotologie.

10 Le plus souvent, ces remarques concernent leur circulation dans « l'argot commun ». Citons des commentaires qui accompagnent nos 50 lexèmes choisis : « *bagnole*, *engueuler* – [sont] passés dans l'usage familier ; *flic* – ce mot très répandu a partiellement perdu son caractère péjoratif ». Nous sommes donc d'avis que si la notion d'*argot commun*, exposée ci-dessus, était communément reconnue et adoptée de façon unifiée en linguistique française, bon nombre d'hésitations sur la coappartenance des mots à plusieurs niveaux / registres pourrait être éliminé.

d'une combinaison des marques, voire même en créant une nouvelle marque. Cela aurait pour but d'éviter la présence de lourds commentaires qu'on retrouve dans ce dictionnaire d'argot aussi bien que dans le *Dictionnaire du français non conventionnel* (Cellard & Rey 1991, désormais DFNC)¹¹ pour la même série de mots¹². Ces annotations nombreuses et peu économiques ne permettent pas au lecteur d'accéder rapidement à l'information univoque, même si, pour les locuteurs non natifs notamment, elles apportent des renseignements importants sur la propagation du mot dans l'usage courant et sur sa fréquence. Cependant, notons que ces annotations apportent également le témoignage de l'hésitation des auteurs quand il s'agit du classement d'un lexème décontextualisé dans un registre précis. De plus, les gloses renvoient souvent à l'expressivité et à la connotation (que nous avons soulignées ci-dessous) :

- dans le DA par exemple : « *bide*, mot *expressif*, resté très usuel » ; « *dingue*, avec cette *valeur emphatique*, l'adjectif est à la mode dans les années 80 » ;
- dans le DFNC, c'est le cas des expressions « *bagnole*, vieilli ou mauvaise automobile, par extension et avec une *nuance de sympathie amusée* : toute automobile » ; « *engueuler*, très large diffusion du mot, aujourd'hui à peine familier, tient à ce que le français conventionnel ne dispose, pour exprimer cette notion, que de verbes faibles ou isolés, d'allure archaïque ou de périphrases *peu expressives* » ; « *gueule*, le mot est neutre plus ou moins, très usuel, est senti comme "*vigoureux*" mais non argotique » ; et enfin : « *pinard*, usuel, mais sans "*coloration*" ».

Les remarques telles que *mot expressif*, *nuance ou valeur emphatique*, *coloration*, etc., sont une preuve que l'expressivité joue un rôle primordial dans ce type de lexique et ceci peut être un signe en faveur de l'intégration, en lexicologie française, de la notion d'expressivité en tant que marque lexicographique.

4. MINI-ANALYSE DIACHRONIQUE ET DIASTRATIQUE DU MARQUAGE DANS LE PR ET DANS LE PL

En tant qu'étrangère, nous nous rendons compte que la lexicographie française manque d'un moyen qui exprimerait l'usage courant du lexème ; mais une telle variabilité des marques métalinguistiques, qui ont été attribuées au lexique issu de l'argot traditionnel (selon le DA), remet en cause, au milieu du désarroi provoqué par les approches disparates, le rôle de l'argot dans le sens moderne du terme qui est, à notre avis, articulé autour de sa fonction primordiale : la fonction expressive. Passons donc maintenant en revue des marques métalinguistiques dans le PR et le PL pour les 50 lexèmes évoqués ci-dessus. Dans les années 1960, les expressions recensées dans le DA comme argotiques (ce qui sous-entend plutôt d'origine argotique), ne sont qu'au nombre de 6 (PL) et 7 (PR) qui sont marquées *arg.* Or, dans les années 1990, elles sont le plus souvent devenues *pop.*, « populaires » (26 dans le PL, soit 79 %, et 37 dans le PR, soit 77 %). Cette constatation ne devrait pas surprendre : la définition de l'abrévia-

11 Malgré un titre très prometteur, les auteurs finissent par classer le lexique selon le modèle traditionnel tripartite : *standard* (ici conventionnel), *familier* (ici semi-conventionnel) et *populaire / argotique / vulgaire* (ici non conventionnel).

12 Citons *bide* – très usuel ; *cinglé* – après avoir été non conventionnel, l'adjectif n'est plus que familier ; *fric* – devenu à peine familier et extrêmement courant.

tion *pop.* dans le PR indique explicitement que le lexique populaire est « souvent argot ancien répandu ».

À cette époque-là, la marque *fam.* n’est quasiment pas utilisée (uniquement pour *pisser* chez *Larousse* et pour *putain* chez *Robert*). Dans les années 1990, la marque *pop.* est progressivement évincée au profit de la marque *fam.* (particulièrement dans le PR, 31 occurrences, moins fréquemment dans le PL, 12 occurrences, ces entrées ne sont pas mises en italique dans le tableau joint en annexe). La marque *pop.* reste le plus souvent inchangée dans le PL (13 occurrences), alors que le PR conserve cette marque plus rarement (seulement 3 occurrences).

En somme, les deux dictionnaires attribuent la même marque pour 28 cas dans les années 1960 et pour 25 cas dans les années 1990 ¹³, mais seulement 17 expressions (soit 34 %) ont la même marque dans les deux dictionnaires des éditions respectives (ces entrées ne sont pas mises en caractères gras dans le tableau en annexe).

Une comparaison de l’évolution des marques, mais également du système d’attribution des marques dans ces deux dictionnaires ¹⁴ peut éclaircir les liens réels et supposés entre les différents registres du sub-standard. Nous proposons de donner un aperçu de ces marques afin d’observer leur fonctionnalité et leur rapport avec l’argot dans le sens moderne du terme, évoqué ci-dessus.

4.1 MARQUE « PÉJORATIF » / « INJURIEUX »

marque <i>injur.</i>	<i>Petit Larousse</i>		<i>Petit Robert</i>	
	1969	1993	1967	1994
pédé	0	vulg. et <i>injur.</i>	pop.	fam.
putain	pop.	vulg. et <i>injur.</i>	fam. et vulg.	<i>péj.</i> et vulg.
pute	0	vulg. et <i>injur.</i>	pop. et vulg.	<i>péj.</i> et vulg.

Cette marque à double appellation – PL préfère l’étiquette « injurieux », PR l’étiquette « péjoratif » – semble très souvent être liée à la marque *vulg.* comme si ces expressions étaient un sous-groupe du registre « vulgaire ». Ici aussi, l’approche adoptée nous permet de constater que la coappartenance des mots à plusieurs registres est évidente ¹⁵. En effet, quel que soit le registre rajouté (*vulg.* ou *fam.*), le caractère injurieux et agressif semble être indéniable pour ce type de « parole offensante » (Guiraud 1991 [1975] : 31).

Le constat d’une banalisation importante pour certaines insultes dans les dictionnaires est alarmant : on se demande pourquoi les deux dictionnaires ne marquent pas de manière identique le lexème *tapette* qui est employé comme une injure, de la même façon que *pédé* (qui reçoit cette marque dans le PL

13 La marque *injur.* du *Larousse* correspond, selon toute évidence, à la marque *péj.* du *Robert*.

14 Le PL n’abandonne pas la marque *pop.* aussi systématiquement que le PR et il s’avère plus puriste que ce dernier.

15 Prenons pour exemple le mot *pédale* qui est marqué *vulg. et injur.* dans le PL, mais *fam. et péj.* dans le PR.

uniquement), *pute* ou *putain*. Un sème inhérent de tous ces termes reste pourtant toujours négatif, méprisant, outrageant.

Ainsi, les marques *péj.* / *injur.* ne renvoient pas à un registre précis. Par son emploi, le type de lexique qui s'y rattache touche à l'intimité de l'autre, aux tabous individuels ou collectifs : dans cette optique, il correspond bien à la définition moderne du lexique argotique dont il fait entièrement partie.

5.2 MARQUE « VULGAIRE »

marque <i>vulg.</i>	<i>Petit Larousse</i>		<i>Petit Robert</i>	
	1969	1993	1967	1994
bite	0	<i>vulg.</i>	0	<i>vulg.</i>
se branler	0	0	0	<i>vulg.</i>
chier	0	<i>vulg.</i>	très <i>vulg.</i>	fam. et <i>vulg.</i>
gonzesse	0	arg.	<i>vulg.</i>	fam.
pédé	0	<i>vulg.</i> et <i>injur.</i>	pop.	fam.
pisser	fam.	très fam.	<i>vulg.</i>	fam.
putain	pop.	<i>vulg.</i> et <i>injur.</i>	fam. et <i>vulg.</i>	péj. et <i>vulg.</i>
pute	0	<i>vulg.</i> et <i>injur.</i>	pop. et <i>vulg.</i>	péj. et <i>vulg.</i>
tapette	0	<i>vulg.</i>	pop. et <i>vulg.</i>	fam. et <i>vulg.</i>

Dans la terminologie lexicographique actuelle, bien que les adjectifs *vulgaire* et *populaire* soient identiques d'un point de vue étymologique, les deux qualificatifs renvoient à deux registres différents. Sous la marque *vulg.*, on trouve aujourd'hui surtout des mots obscènes ou des mots grossiers référant au sexe et à la scatologie. Nous insistons sur le mot « aujourd'hui », puisque la notion de vulgarité ne commence à se réduire à la transgression des tabous culturels que depuis quelques dizaines d'années. Nous pouvons observer sur notre échantillon que la marque *vulg.* a subi un changement référentiel considérable ¹⁶.

D'un point de vue sémantique, la vulgarité se présente en tant que valeur soit dénotative, soit connotative :

Un terme connoté « vulgaire » a tendance à vulgariser, par contagion, le signifié, donc le dénoté auquel il renvoie ; inversement, les termes stylistiquement « normaux » qui désignent des réalités sexuelles ou scatologiques ont tendance à être perçus comme « bas » dans la mesure où la dévalorisation qui s'attache au contenu finit par déteindre sur le signifiant. (Kerbrat-Orecchioni 1999 : 84)

La linguistique tchèque, conformément à la tradition anglo-saxonne, désigne ce type de lexique comme faisant partie des vulgarismes, notion peu utilisée en linguistique française. Cependant, cette notion nous semble pertinente compte

¹⁶ En comparaison au PR 1967, celui de 1994 spécifie que c'est le domaine des tabous qui est privilégié et que le type de discours est bien plus déterminant que le fait de savoir si la personne est bien élevée ou non. Mieux encore, le PL spécifie que les tabous sont « le plus souvent d'ordre sexuel ou excrémental » (PL 1993 : 24).

tenu de l'absence de connotations qui s'y rattachent et de son appartenance logique au paradigme des *-ismes* (argotisme, slanguisme, etc.), paradigme qui permet une certaine économie.

Les lexicographes ont longtemps pratiqué la tradition normative qui ignorait délibérément tous les vulgarismes. Paradoxalement, ce sont les dictionnaires d'argot qui, les premiers, ont recensé les gros mots, les injures et les jurons, connus de tous et employés souvent depuis des siècles. Au sens strict du terme, ces mots n'ont que peu de traits argotiques (au sens traditionnel du mot), mais c'est justement parce qu'ils sont ignorés par les lexicographes des dictionnaires officiels qu'ils sont répertoriés par ceux de la langue verte. Or, ce qui nous semble être proche de l'argot, c'est le caractère cryptique de nombreuses métaphores référant aux sujets jugés vulgaires (elles sont cryptiques au moins au moment de leur création avant d'être banalisées). Dans l'optique d'une définition moderne de l'argot, les vulgarismes sont inclus dans ce dernier, car ils ont un caractère transgressif par rapport à la norme, à la convention. Ils possèdent également une forte valeur expressive inhérente qui impressionne l'interlocuteur et qui peut servir soit comme signe de connivence, soit comme signe d'hostilité.

Même si l'on constate une importante « libéralisation » des marques, n'oublions pas de rappeler le revers de la médaille : les emplois populaires ou familiers des mots dont le sens primaire est vulgaire ne sont plus pris en compte et l'évaluation de ce type de mots devient plus sévère¹⁷. Or, faut-il privilégier le critère de la bienséance, issu de la norme conventionnelle désuète, qui exclut tous les actes frappés d'un tabou qui sont renvoyés au registre vulgaire ou bien le critère pragmatique qui observe l'usage courant de ces verbes dans le discours familier où ils sont utilisés le plus souvent dans des locutions figées qui ignorent le sens primaire tabouisé et qui, par la suite, atténuent l'expressivité négative de ces verbes mêmes ? Ces deux approches s'affrontent sans cesse. Dans notre échantillon, on voit bien que le PR résout ce problème en combinant les marques.

4.3 MARQUE « POPULAIRE »

À l'époque actuelle de perméabilité linguistique entre les classes sociales, la notion de « français populaire » semble être un stéréotype social particulièrement stigmatisant, issu de l'opposition puriste entre « bien parler » et « mal parler », entre un français « cultivé » et un français « vulgaire » (Gadet 1992 : 26-27). Cette notion n'est pourtant pas facile à contourner, car la marque « populaire » ne cesse d'être employée dans les dictionnaires de langue française. Or, alors que la notion de français populaire est tout à fait défendable au niveau phonique (Gadet 1992), les tentatives pour l'étendre à d'autres plans linguistiques, notamment au plan lexical, se heurtent à des difficultés de délimitation de ce qui est propre à ce registre de langue (v. Abecassis 2003). Si l'on ne tient pas compte des catégorisations au fondement purement idéologique, on arrive à la conclusion que c'est notamment le registre du français familier et celui du français populaire qui se recouvrent considérablement (sous l'angle diaphasie / diastatie).

¹⁷ Dans le PR, par exemple, *putain* cesse d'être considéré comme *fam.* et *vulg.* au profit de *péj.* et *vulg.* ou *pute* n'est plus *pop.* et *vulg.*, mais strictement *péj.* et *vulg.*, etc.

marque <i>pop.</i>	<i>Petit Larousse</i>		<i>Petit Robert</i>	
	1969	1993	1967	1994
becter / bequeter	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
bol	0	fam.	<i>pop.</i>	fam.
dingue	0	fam.	<i>pop.</i>	fam.
fric	arg.	fam.	<i>pop.</i>	fam.
fringues	arg.	fam.	<i>pop.</i>	fam.
froc	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
godasse	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
gonzesse	0	arg.	vulg.	fam.
grolle	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
gueule	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
mec	0	fam.	<i>pop.</i>	fam.
pédé	0	vulg. et injur.	<i>pop.</i>	fam.
pinard	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
pognon	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
pompe	0	fam.	<i>pop.</i>	<i>pop.</i>
portugaise	0	<i>pop.</i>	<i>pop.</i>	arg. fam.
putain	<i>pop.</i>	vulg. et injur.	fam. et vulg.	péj. et vulg.
pute	0	vulg. et injur.	<i>pop.</i> et vulg.	péj. et vulg.
robert(s)	0	<i>pop.</i>	<i>pop.</i>	fam.
se saper	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
se tailler	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>
tapette	0	vulg.	<i>pop.</i> et vulg.	fam. et vulg.
tarin	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
taule	<i>pop.</i>	<i>pop.</i>	arg.	arg.
tif(s)	0	<i>pop.</i>	<i>pop.</i>	fam.
se tirer	0	<i>pop.</i>	<i>pop.</i>	fam.
tronche	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	fam.
turbine	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i> et vieilli

Remarque : Afin de rester concise, je n'ai pas intégré dans ce tableau les 12 lexèmes qui ont connu, dans les deux dictionnaires, la même évolution de *pop.* vers *fam.* (ils sont répertoriés sans italique dans le tableau final).

La notion de français populaire est donc à éviter ; elle est connotée différemment en linguistique et auprès des non-spécialistes, et redondante en ce qui concerne les marques lexicographiques. L'examen de nos lexèmes montre que la marque « populaire » est la plus arbitraire de toutes, vu que son flottement est énorme non seulement d'un point de vue diachronique, mais également en synchronie¹⁸. Le « populaire » se présente alors comme un fossile qui n'a plus aucune fonction et qui est facilement remplaçable par d'autres marques d'usage.

Finalement, le lien entre le populaire et l'argotique est plus important que ce qu'on peut penser si l'on s'en tient uniquement aux marques lexicographiques. Tandis que dans notre mini-corpus, aucun mot n'a évolué de la marque « argotique » vers la marque « populaire », dans la réalité, le passage entre ces deux catégories a été très fréquent (v. Červenková 2002), mais pour des raisons idéologiques, les deux marques ont été conservées. Selon nos observations, c'est surtout le critère de la vulgarisation du lexème qui s'avère être le plus saillant quant à la démarcation entre le populaire et l'argotique.

4.4 MARQUE « FAMILIER »

La marque *fam.* joue aujourd'hui un rôle très important¹⁹ puisqu'elle recouvre tout le niveau intermédiaire entre le standard et les formes soit connotées socialement (populaire, argotique), soit tabouisées puisqu'elles vont contre la bienséance (vulgaire, injurieux). Le français familier est surtout un style conversationnel (l'expression est fréquente en tchèque – elle correspond à l'expression anglaise « colloquial speech »), une variation diaphasique propre à tout locuteur, quelle que soit son origine sociale. Si l'on observe notre mini-corpus, les 50 mots cités figurent comme argotiques dans le *Dictionnaire de l'argot*, mais 20 dans le PL et 37 dans le PR portent pourtant la marque « familier ». Ceci est une preuve de l'existence de l'« argot commun », notion qui correspond plus ou moins au *slang* anglais, évoqué au début de cet article.

La notion d'argot commun permet de mettre en évidence à la fois l'origine du lexème (issu de la pratique argotique) et son usage fréquent, sa vulgarisation (commun, devenu familier). C'est également dans cette optique que le PR commence à utiliser la marque *arg. fam.*, c'est-à-dire la marque désignant « les mots d'argot passés dans le langage familier », selon la définition donnée par les auteurs de ce dictionnaire (la marque ARG. FAM. est mise en majuscules dans le tableau ci-après). L'expression *brique*, portant la marque *arg.* en 1967 devient ainsi *arg. fam.* en 1994. En revanche, l'expression *portugaise* reçoit également cette marque en 1994 alors qu'en 1967, elle était désignée comme *pop.* et non comme *arg.* L'idée des auteurs est bonne, mais cette tentative se heurte à un manque de fermeté : pourquoi les expressions *casse* ou *se casser*, par exemple, passent-elles directement du registre argotique au registre familier, sans recevoir la marque *arg. fam.* ?

En somme, la perméabilité lexicale de l'argot des groupes vers l'argot commun est un des principaux axes de l'innovation lexicale.

18 Dans notre mini-corpus, il n'y a que deux expressions, *se tailler* et *turbin*, qui sont marquées par les deux dictionnaires comme *pop.* Dans les autres cas, l'attribution de la marque *populaire* diverge dans les éditions des années 1990. Le PR l'évince au profit de la marque *fam.* (31 lexèmes) et la garde uniquement pour 3 entrées sur 50 : *pompe*, *se tailler* et *turbin*. Le PL est plus fidèle à cette marque controversée (on répertorie ainsi 17 lexèmes sur 50).

19 Quasiment non utilisée dans les années 60, cette marque doit son énorme succès à la connotation stigmatisante de la marque « populaire » qui lui cède la place pour désigner, de façon tout à fait neutre, des lexèmes employés couramment dans une conversation un peu libre, dans la langue parlée relâchée.

marque <i>fam.</i>	<i>Petit Larousse</i>		<i>Petit Robert</i>	
	1969	1993	1965	1994
bol	0	<i>fam.</i>	pop.	<i>fam.</i>
brique	arg.	<i>fam.</i>	arg.	ARG. FAM.
casse	arg.	arg.	arg.	<i>fam.</i>
se casser	0	arg.	arg.	<i>fam.</i>
chier	0	vulg.	très vulg.	<i>fam.</i> et vulg.
dingue	0	<i>fam.</i>	pop.	<i>fam.</i>
fric	arg.	<i>fam.</i>	pop.	<i>fam.</i>
fringues	arg.	<i>fam.</i>	pop.	<i>fam.</i>
gonzesse	0	arg.	vulg.	<i>fam.</i>
mec	0	<i>fam.</i>	pop.	<i>fam.</i>
pédé	0	vulg. et injur.	pop.	<i>fam.</i>
pisser	<i>fam.</i>	très <i>fam.</i>	vulg.	<i>fam.</i>
pompe	0	<i>fam.</i>	pop.	pop.
portugaise	0	pop.	pop.	ARG. FAM.
putain	pop.	vulg. et injur.	fam. et vulg.	péj. et vulg.
tapette	0	vulg.	pop. et vulg.	<i>fam.</i> et vulg.

Remarque : comme dans le tableau précédent, nous écartons de ce tableau les évolutions de *pop.* vers *fam.* dans les deux dictionnaires (12 lexèmes) ainsi que 13 autres lexèmes qui sont marqués *pop.* > *fam.* dans le PR et *pop.* > *pop.* (10 lexèmes) ou néant > *pop.* (3 lexèmes) dans le PL.

4.5 RETOUR SUR LA MARQUE « ARGOTIQUE »

Finalement, si l'on examine les expressions portant la marque *arg.* dans le mini-corpus appliqué aux deux dictionnaires observés, on ne voit pas bien quels sont ses critères d'attribution. Dans le PL de 1969, 6 expressions la portent, à savoir *brique*, *casse*, *esgourde*, *fric*, *fringues* et *tire*. Tandis que *brique*, *fric* et *fringues* sont passés dans le registre familier en 1993, *casse*, *esgourde* et *tire* portent toujours la marque « argotique » et *gonzesse*, *se casser* et *châsse*, nouvellement entrés dans le dictionnaire, ont directement reçu la marque *arg.* Pour le PR de 1967, *brique*, *casse*, *esgourde* et *tire* sont argotiques, ainsi que *taule* (qui, dans le PL, est toujours populaire), *châsse* et *se casser*. Or, en 1994, alors que *brique* obtient la marque *arg.fam.*, *casse* et *se casser* deviennent sans raison simplement familiers, et *châsse*, *esgourde*, *taule* et *tire* gardent leur marque argotique.

Nous énumérons ces différents classements pour montrer non seulement à quel point cette marque est vague (les deux dictionnaires ne s'accordent sur la marque argotique que dans 3 expressions), mais surtout pour montrer qu'il s'agit là d'expressions-clichés pour évoquer le vieil argot, l'argot vieilli, la fameuse « langue verte » qui est, à cause de sa notoriété, connue par un large public et qui est donc très peu cryptique.

marque <i>arg.</i>	<i>Petit Larousse</i>		<i>Petit Robert</i>	
	1969	1993	1965	1994
brique	<i>arg.</i>	fam.	<i>arg.</i>	ARG. FAM.
casse	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>	fam.
se casser	0	<i>arg.</i>	<i>arg.</i>	fam.
châsse	0	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>
esgourde	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>
fric	<i>arg.</i>	fam.	pop.	fam.
fringues	<i>arg.</i>	fam.	pop.	fam.
gonzesse	0	<i>arg.</i>	vulg.	fam.
portugaise	0	pop.	pop.	ARG. FAM.
taule	pop.	pop.	<i>arg.</i>	<i>arg.</i>
tire	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>

CONCLUSION

Ayant observé les chevauchements entre le vulgaire et l'argotique d'une part, l'argotique, le populaire et le familier d'autre part, mais aussi les doubles emplois possibles des mots familiers au contenu vulgaire, on comprend vite pourquoi toutes les expressions de notre mini-corpus peuvent facilement figurer dans un dictionnaire d'argot. Tout se passe comme s'il y avait des zones de passage allant dans le sens :

(vulgaire) → ARGOTIQUE → populaire → familier

Malgré toutes les tendances observées ci-dessus consistant à remplacer les marques *populaire* et *argotique* par *familier*, excepté pour le cas des termes hors d'usage courant qui devraient garder la marque *pop.* ou *arg.* avec la mention *vieilli*, nous pensons qu'une grande part des francophones ne va pas se priver aussi facilement du marquage de la « connotation sociale » des mots en rangeant la totalité des termes sous l'étiquette *familier*. La nouvelle marque combinée, *arg. fam.* du PR, est prometteuse car elle permet de repérer à la fois l'origine sociale et l'emploi stylistique. Cette nouvelle marque est de plus en plus fréquemment utilisée, ce qui permet de justifier l'existence d'un argot commun (ainsi que le droit de cité de dictionnaires d'argot comme celui de Larousse, etc.). Même si elle ne nous paraît pas aussi efficace que la marque *expr. tchèque*, c'est une tendance qui permet d'éviter les longs commentaires sur la coappar-

tenance d'un mot aux différents registres que nous avons évoqués. En somme, lier le social au situationnel est un compromis entre l'approche traditionnelle qui respecte l'étymologie et l'approche moderne basée sur la pragmatique du sens.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ABECASSIS Michaël, 2003, « Le français populaire : *a valid concept?* », *Marges linguistiques*, 6, p. 116-132.
- BALLY Charles, 1921, *Traité de stylistique française* (2 vol.), Heidelberg, Carl Winter (1^{re} éd. Paris, Klincksieck, 1909).
- BALLY Charles, 1935, *Le Langage et la Vie*, Zurich, Max Niehans (1^{re} éd. Genève, Atar, 1913).
- BLANCHE-BENVENISTE Claire, 2000, *Approches de la langue parlée en français*, Paris, Ophrys.
- CELLARD Jacques et REY Alain, 1991, *Dictionnaire du français non conventionnel*, Paris, Hachette (1^{re} éd. Paris, Masson, 1980).
- COLIN Jean-Paul, MÉVEL Jean-Pierre et LECLÈRE Christian, 1990, *Dictionnaire de l'argot*, Paris, Larousse.
- CORBIN Pierre, 1989, « Les marques stylistiques / diastratiques dans le dictionnaire monolingue », dans F. Hausmann *et alii*, *Wörterbücher / Dictionaries / Dictionnaires*, t. 1, Berlin und New York, Walter de Gruyter, p. 937-946.
- ČERVENKOVÁ Marie, 2002, *L'Enrichissement du français standard des sources argotiques*, Thèse sous la direction de Růžena Ostrá, Brno, Université Masaryk de Brno.
- FRANÇOIS-GEIGER Denise, 1989, *L'Argoterie, recueil d'articles*, Paris, Sorbonnargot.
- FRANÇOIS-GEIGER Denise, 1991, « Panorama des argots contemporains », dans D. François-Geiger et J.-P. Goudaillier (éds), *Parlures argotiques, Langue française*, 90, p. 5-9.
- GADET Françoise, 1992, *Le Français populaire*, Paris, Puf, « Que sais-je ? », n° 1172.
- GOUDAILLIER Jean-Pierre, 2002, « De l'argot traditionnel au français contemporain des cités », dans J.-P. Goudaillier (éd.), *Argots et Argotologie, La Linguistique*, 38 (1), p. 5-24.
- GUIRAUD Pierre, [1975] 1991, *Les Gros Mots*, Paris, Puf, « Que sais-je ? », n° 1597.
- JAKLOVÁ Alena, 1999, « Budeme argot nově definovat », *Slovo a slovesnost*, 60, p. 293-300.
- JOLIN-BERTOCCHI Sophie, 2003, *Les Niveaux de langage*, Paris, Hachette.
- KERBRAT-ORECCHIONI Catherine, 1999, *L'Énonciation*, Paris, Armand Colin.
- LODGE Anthony, 1997, « The pragmatics of slang », *Web Journal of Modern Language Linguistics*, <http://wjml.ncl.ac.uk/issue02/lodge/htm>.
- MACHEK Václav, 1930, *Studie o výrazech expresivních*, Praha, FF Univerzity Karlovy.
- PODHORNÁ-POLICKÁ Alena, 2007, *Peut-on parler d'un argot des jeunes? Analyse lexicale des universaux argotiques du parler de jeunes en lycées professionnels en France (Paris, Yzeure) et en République tchèque (Brno)*, thèse

en cotutelle sous la direction de Jean-Pierre Goudaillier et Marie Krčmová, Université René Descartes, Paris, et Université Masaryk, Brno.

RASTIER François, 1987, *Sémantique interprétative*, Paris, Puf.

SOURDOT Marc, 1991, « Argot, jargon, jargot », dans D. François-Geiger et J.-P. Goudaillier (éds.), *Parlures argotiques, Langue française*, 90, p. 13-27.

SOURDOT Marc, 2002, « L'argotologie : Entre forme et fonction », dans J.-P. Goudaillier (éd.), *Argots et Argotologie, La Linguistique*, 38 (1), p. 25-39.

WALTER Henriette, 2005, *Le français dans tous les sens*, Paris, Le Livre de Poche, (1^{re} éd. 1988).

ZIMA Jaroslav, 1963, *Expresivita slova v současné češtině. Studie lexikologická a stylistická*, Praha, Nakladatelství Československé akademie věd.

ABRÉVIATIONS DES DICTIONNAIRES

DA = *Dictionnaire d'argot* 1990

DFNC = *Dictionnaire du français non conventionnel* 1991

PR = *Petit Robert* 1967 et 1994

PL = *Petit Larousse* 1969 et 1993

ANNEXE

Tableau de la dynamique des changements des marques d'usage dans le *Petit Larousse* et dans le *Petit Robert*.

expression « marquée »	<i>Petit Larousse</i>		<i>Petit Robert</i>	
	1969	1993	1967	1994
bagnole	pop.	fam.	pop.	fam.
se barrer	pop.	fam.	pop.	fam.
becter / bequeter	pop.	pop.	pop.	fam.
bide	pop.	fam.	pop.	fam.
bite	0	vulg.	0	vulg.
bol	0	fam.	pop.	fam.
bossier	pop.	fam.	pop.	fam.
bouffer	pop.	fam.	pop.	fam.
se branler	0	0	0	vulg.
brique	arg.	fam.	arg.	arg. fam.
caïd	pop.	fam.	pop.	fam.
casse	arg.	arg.	arg.	fam.
se casser	0	arg.	arg.	fam.
châsse	0	arg.	arg.	arg.
chier	0	vulg.	très vulg.	fam. et vulg.
cinglé	pop.	fam.	pop.	fam.

<i>dingue</i>	<i>0</i> ²⁰	<i>fam.</i>	<i>pop.</i>	<i>fam.</i>
<i>engueuler</i>	<i>pop.</i>	<i>fam.</i>	<i>pop.</i>	<i>fam.</i>
<i>esgourde</i>	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>
<i>flic</i>	<i>pop.</i>	<i>fam.</i>	<i>pop.</i>	<i>fam.</i>
<i>frangin, -e</i>	<i>pop.</i>	<i>fam.</i>	<i>pop.</i>	<i>fam.</i>
<i>fric</i>	<i>arg.</i>	<i>fam.</i>	<i>pop.</i>	<i>fam.</i>
<i>fringues</i>	<i>arg.</i>	<i>fam.</i>	<i>pop.</i>	<i>fam.</i>
<i>froc</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>godasse</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>gonzesse</i>	<i>0</i>	<i>arg.</i>	<i>vulg.</i>	<i>fam.</i>
<i>grolle</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>gueule</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>mec</i>	<i>0</i>	<i>fam.</i>	<i>pop.</i>	<i>fam.</i>
<i>oseille</i>	<i>pop.</i>	<i>fam.</i>	<i>pop.</i>	<i>fam.</i>
<i>pédé</i>	<i>0</i>	<i>vulg. et injur.</i>	<i>pop.</i>	<i>fam.</i>
<i>pif</i>	<i>pop.</i>	<i>fam.</i>	<i>pop.</i>	<i>fam.</i>
<i>pinard</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>pisser</i>	<i>fam.</i>	<i>très fam.</i>	<i>vulg.</i>	<i>fam.</i>
<i>pognon</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>pompe</i>	<i>0</i>	<i>fam.</i>	<i>pop.</i>	<i>pop.</i>
<i>portugaise</i>	<i>0</i>	<i>pop.</i>	<i>pop.</i>	<i>arg. fam.</i>
<i>putain</i>	<i>pop.</i>	<i>vulg. et injur.</i>	<i>fam. et vulg.</i>	<i>péj. et vulg.</i>
<i>pute</i>	<i>0</i>	<i>vulg. et injur.</i>	<i>pop. et vulg.</i>	<i>péj. et vulg.</i>
<i>robert(s)</i>	<i>0</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>se saper</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>se tailler</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>
<i>tapette</i>	<i>0</i>	<i>vulg.</i>	<i>pop. et vulg.</i>	<i>fam. et vulg.</i>
<i>tarin</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>taule</i>	<i>pop.</i>	<i>pop.</i>	<i>arg.</i>	<i>arg.</i>
<i>tif(s)</i>	<i>0</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>tire</i>	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>	<i>arg.</i>
<i>se tirer</i>	<i>0</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>tronche</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>fam.</i>
<i>turbin</i>	<i>pop.</i>	<i>pop.</i>	<i>pop.</i>	<i>pop. et vieilli</i>

20 L'entrée *dingue* est manquante, mais le dictionnaire présente une forme *dingo* = fou avec une mention argotique.

DU MÉCANISME CACHÉ
SOUS LA MARQUE LEXICOGRAPHIQUE *FAM.* :
LE CAS DES NOMS

Jackie SCHÖN

Laboratoire Jacques Lordat, Université de Toulouse le Mirail

Résumé — La *familiarité* langagière ressortit à l'*effet* produit lors de l'*emploi* des expressions. Il s'agit donc d'un phénomène *discursif* mais suffisamment codé et stable pour être répertorié dans nos dictionnaires de langue sous les traits de la *marque*. Parmi les éléments éventuellement marqués, seules les *acceptions* familières de noms sont abordées ici. Si l'*acception* n'est que le résultat du processus de signification d'une expression quelconque, lorsque cette acception apparaît marquée, elle désigne – pour une forme donnée – un nouveau sens qui se rajoute à son sens premier. Soit, en français, les lexèmes *andouille* ou *cruche* dont les acceptions secondes, marquées, sont respectivement illustrées dans le *Petit Robert* (éd. 1973) par les séquences : *Sacrée andouille ! Quelle cruche !* Le sens qui leur est, communément, attribué devient : Niais, imbécile... dès lors que le référent désigné par ces termes est une personne, condition que les dictionnaires impliquent sans l'explicitier. Le mécanisme à l'œuvre se révèle à l'évidence : la familiarité naît d'une *incongruité*, celle qui s'obtient par le recouvrement *systématique* d'une classe de dénominatifs des personnes par des classes sémantiques de noms impropres à leur dénomination (noms de comestibles, d'animaux, d'objets domestiques, etc.). Ainsi s'exprime la *subjectivité* du locuteur à l'égard de la personne qu'il désigne.

Mots clés — acception, familier, marque, incongruité sémantique, subjectivité.

Abstract — Familiarity in language relates to the effect produced when expressions are used. It is therefore a discursive phenomenon, coded and stable enough to be listed in language dictionaries under the form of a label. The “familiar” label is examined in this article in order to explain which semantic process underlies this very label. The author takes the examples of two words : *andouille* (basic meaning ‘regional sausage’) and *cruche* (basic meaning ‘pitcher’). Their secondary meanings are labelled in *Le Petit Robert* (1973) as respectively : *Sacrée andouille !* (‘damned idiot!’) and *Quelle cruche !* (‘what a twit!’). The meaning that is commonly attributed to both these words is then: ‘silly, stupid’ – but only if the referent designated by them is a person, a condition which in the dictionaries is implicit rather than made explicit. The mechanism at work here is quite obvious : familiarity is born from incongruity. In our examples the incon-

gruity is the consequence of systematically covering a class of nouns designating people with semantic classes of nouns improperly attributed to them (nouns referring to foods, animals, household objects, etc.). Some *specific* meanings result from a semantic process involving the expression of the speakers' subjectivity towards the people they refer to. This process is then comprised under the *broad* label "familiar usage" concealing the very diverse categories hidden under such a general label.

Keywords — specific meaning, "familiar" label, semantic incongruity, subjectivity.

Il est au moins deux points de départ possibles pour cette étude : soit on part des marques d'usage figurant dans les dictionnaires comme d'un fait établi, un donné non remis en cause, l'objet d'étude étant, alors, le contenu déjà élaboré des dictionnaires ; soit on prend appui sur l'observation des régularités linguistiques pour aboutir à l'introduction des marques par le lexicographe.

Jusqu'à présent, j'avais toujours adopté le second des points de vue, me situant en quelque sorte en amont du phénomène familier dont le signalement dictionnaire ne constituait qu'une conclusion prévisible. Compte tenu de ce qui fonde l'unité du thème proposé ici, je me rangerai volontiers en aval du phénomène, au niveau de l'indice *fam.* consigné par les dictionnaires de langue les plus courants, mais ce ne sera jamais que pour mieux remonter jusqu'à ce que cet indice recouvre, pour démontrer ce qu'il dissimule.

Pourquoi ne pas interroger d'abord la notion même de *familiarité*, appréhender ses retombées dans le champ du social pour, enfin et surtout, cerner les caractéristiques du fait *linguistique* familier ?

Étymologiquement, les termes « familiarité » et « familier » (du lat., respectivement *familiaritas* et *familiaris*) concernent, certes, la famille mais, ce qui est tout aussi important, ils appellent la notion d'*effet* causé / ressenti lors d'un contact. Il n'est rien d'intrinsèquement *familier* ; tout être, lieu, événement ou objet ne saurait être *familier* que relativement à quelqu'un ou *pour* quelqu'un au cours d'une mise en relation. En l'occurrence, il s'agira d'un échange verbal ; nous sommes au cœur du phénomène discursif.

Parmi les composantes de cet *effet familier* notons les thèmes du retour, de la remémoration, de la *remembrance*, de la re-connaissance d'un déjà connu, vu, vécu. Dans son bel article traduit en français sous le titre « L'inquiétante étrangeté », Freud (1919) montre que la résurgence d'éléments assignés au domaine *privé* de la famille et voués à y demeurer manifeste le retour à un stade infantile refoulé, atteste d'une régression vers un stade psychique des plus archaïques, jusqu'au refuge maternel dont le petit d'homme porterait à jamais la nostalgie.

Un échange de type *familier* entraîne un rapprochement entre les interlocuteurs, les amenant à déroger ensemble à des règles de bienséance – ou posées comme telles par les groupes dominants –, appelons-les plutôt : règles de neutralité langagière. Il n'en reste pas moins que la pratique de la familiarité langagière met son utilisateur en situation de vulnérabilité dangereuse. Ainsi, en usant familièrement de la langue, le locuteur se montre-t-il agressif et narcis-

sique à l'instar du tout petit enfant (voir aussi Schön 1999), mais aussi démunie et dépendant que lui, tandis que ce que prône le code des convenances est, précisément, le respect de la distance optimale entre chacun des protagonistes, le « chacun à sa place » dans la sphère des jeux de rôles sociaux.

Ces considérations ne nous ont pas éloignés de notre objectif premier qui reste le dégagement des spécificités du langage qualifié de *familier*. L'indice qui, dans les dictionnaires, en matérialise la stigmatisation sociale vaut avertissement : l'emploi de telle expression (lexème, syntème, syntagme ou tournure syntaxique) ne contrevient pas à la langue mais égratigne le code des convenances qui en régit la gestion sociale. Cet emploi ne convient pas à n'importe quelle situation de communication, une forme marquée marque son utilisateur et quand on *marque*, on *marque mal*, forcément !

Après ce préambule, utile pour fournir à ce qui suit un cadre élargi, reprenons notre réflexion sur ce qui se cache sous la marque *fam*.

Assez naïvement, j'avais longtemps considéré le dictionnaire comme une tentative de consignation d'un certain état de langue momentanément suspendu mais en revenant, notamment, à certains écrits d'Alain Rey, j'ai dû reconnaître qu'il n'avait pas tort en constatant que « la plupart des linguistes vivent souvent sur l'idée fausse qu'un dictionnaire décrit une langue, un système » (Rey 1995 : 108).

Du même coup – et puisque j'étais de ceux-là –, la conception que je me faisais de l'introduction des exemples marqués *fam*. dans le dictionnaire s'est trouvée mise à mal. En effet, tant que je voyais dans le dictionnaire une concrétisation du système abstrait Langue (au sens saussurien du terme), l'introduction d'éléments qualifiés de *familiers* m'apparaissait comme l'entrée, quasiment en force, de fragments de discours dans la langue, lesquels assuraient aussi son enrichissement et son renouvellement. En revanche, si je devais convenir de ce que le dictionnaire était constitué de discours, il m'incombait de dégager la spécificité de ceux qui accompagnent les éléments marqués dans le lexique.

Une précision de méthode s'impose : l'indice *fam*. s'applique soit à des entrées lexicales monématiques ou synthématiques, soit à des sous-entrées concernant des *emplois particuliers* de ces unités, délivrant alors des *acceptions*. Une *acception* résulte d'un rapport entre une forme et un ou des sens qui naissent sous certaines *conditions de ses emplois*. Pour l'heure, seul ce dernier phénomène nous intéresse et nous nous bornerons à l'examen des *acceptions familiales* de lexèmes, les unités linguistiques plus étendues ou complexes exigeant des approches adaptées (voir pour exemple Schön 1997). Enfin, parmi les lexèmes, seuls seront pris en compte les lexèmes nominaux dans la mesure où les lexèmes verbaux ont un comportement linguistique différent (Schön 2002).

Au sujet des *acceptions*, le fait le plus remarquable est que, pour une unité lexicale donnée, au moins un sens nouveau vient s'ajouter à celui ou ceux que lui attribue le dictionnaire sans mention d'aucune contrainte d'emploi. Une polysémie particulière se crée ainsi qui peut provoquer ambiguïté d'interprétation des énoncés mais sans effacement des contenus sémantiques des lexèmes que nous appellerons *sens premiers*, ne serait-ce que par commodité. Prenons en exemple des lexèmes banals que certains de leurs usages dotent d'au moins un doublon familier, soit : *cruche*, *tarte*, *andouille*, *courge*, *navet*, *poire*, *nouille*, *cornichon* et, pour certains d'entre eux, observons de plus près leurs marques et

acceptions respectives telles qu'elles apparaissent dans un échantillon restreint à trois dictionnaires dont deux éditions d'un même auteur espacées de cinq ans¹ :

- *andouille* :

Pop. : Niais, imbécile (PR 73) ;

Pop. : Niais, imbécile « *Il ronchonna : sacrée andouille ! Pas fichu seulement de fabriquer une veste !* » (Courteline) (PR 78).

Fam. : Personne sotte ou maladroite. *Espèce d'andouille !* (PL).

- *tarte* :

Pop. : Coup, gifle. *Adj. Fam.* : Laid, sot et ridicule, peu dégourdi, *V. cloche.* (PR 73)

Pop. : Coup, gifle. *Adj. (Arg. v. 1900) Fam.* : Laid, sot et ridicule, peu dégourdi. (PR 78)

Fam. : Gifle. *Adj. Fam.* : Stupide, ridicule, insignifiant. *Garçon tarte. Film tarte.* (PL)

- *cruche* :

Fam. : Personne niaise, bête et ignorante. *C'est une vraie cruche ; Quelle cruche ! V. Imbécile.* (PR 73 & 78)

Fam. : Personne niaise, stupide. (PL)

- *courge* :

Pop. : Imbécile, « gourde ». (PR 73 & 78)

Aucune mention dans le PL.

La remarque première – qui n'est pas forcément la plus importante – porte sur la variabilité des appréciations dictionnaires. D'un dictionnaire à l'autre, d'une édition à une autre, on peut trouver, pour une même expression, des notations diverses : *arg.*, *pop.*, *fam.*, *vulg.*, ou autre encore avec, cependant, un point commun entre toutes, elles indiquent invariablement une stigmatisation sociale de l'usage. Ne voyant dans ces fluctuations de marquage qu'un épiphénomène, nous en remettons l'examen pour nous focaliser uniquement sur ce qui nous paraît l'essentiel, à savoir la description du mécanisme à l'œuvre lors du processus de *familiarisation*. J'entends par *familiarisation* le processus selon lequel des emplois discursifs sporadiques se répandent dans la communauté linguistique pour finir par s'installer dans les usages courants et être suffisamment stabilisés et maîtrisés par l'ensemble des locuteurs pour figurer dans les dictionnaires².

Quelle est la structure syntaxique des discours qui introduisent les emplois marqués des lexèmes ? Ils se présentent sous la forme d'énoncés complets minima, des énoncés-phrases, c'est-à-dire des énoncés dans lesquels s'effectue, obligatoirement, une prédication. Dans les dictionnaires, l'introduction d'emplois familiers s'effectue préférentiellement à l'aide de ce type d'énoncés – souvent des citations – accompagnés d'une ponctuation exclamative. La plupart

¹ Les *Petit Robert*, *Dictionnaire de la langue française*, 1973 et 1978 (PR) et le *Petit Larousse illustré*, 1998 (PL).

² J'avoue ne pas réussir à utiliser la distinction opérée par P. Cadiot et F. Nemo, qui posent dans leur article qu'« Un emploi devient un usage quand il devient dénominatif, quand il désigne une réalité dont il est aussi le nom » (Cadiot & Nemo 1997 : 28).

des prédicats nominaux y sont mis en fonction *attributive* à un objet désigné. En la circonstance et pour un maximum de cohérence, seront privilégiées des désignations de personnes.

Du point de vue sémantique, il est à observer que les significations attachées aux acceptions familières de lexèmes partagent les mêmes caractères stéréotypés : elles sont uniformément imprécises et d'une navrante pauvreté, méprisantes et désagréables à l'égard du référent désigné. Pour ce qui est du flou, du vague et de l'approximation des expressions, que chacun d'entre nous demande à son entourage où passe la distinction entre : *Quelle courge* ou *cruche cette fille !* et *Quelle tarte, cette fille !*

Les hypocoristiques obéissent au même schéma, par ex. : *Quel chou ! ; Qu'il est trognon !* Ils sont souvent précédés d'un possessif qui remplace, en partie, l'exclamation comme on l'entend dans *ma puce, mon canard, ma biche, mon poulet*. Insultes et mots d'amour fonctionnent à l'identique avec, peut-être, un champ lexical plus développé pour les premières que pour les seconds et, bien entendu, la substitution de la sympathie au mépris et d'une possessivité tendre à l'intention désobligeante mais, dans les deux cas, s'opère une réification de l'objet, l'autre est considéré comme une chose, un animal ou un légume.

Quoi de surprenant dans ce constat dès lors qu'on réalise que de tels attributs ne recouvrent aucunement des concepts sinon des *affects* ? Ce sont les affects du locuteur qui s'expriment par ce type de discours et, en matière d'affects, la palette des choix est étroite, bornée aux deux extrêmes par l'euphorie vs la dysphorie : le locuteur aime (et il incorpore) ou il n'aime pas (et il rejette) (v. Schön 2001).

Une fois isolées les différentes pièces d'un assemblage complexe, il reste à en décrire la mise en marche. L'*effet familial* d'un usage provient, comme on pouvait s'y attendre, d'une transgression du code de la langue. En l'occurrence, la transgression repose sur la superposition, opérée *dans* et *par le discours*, de deux classes de noms sémantiquement disjointes.

Ainsi, dans la mesure où nos exemples *familiers* cités s'appliquent, majoritairement, à des personnes, leur regroupement devrait constituer un domaine lexical de « dénominations de la personne ». Or, loin d'entrer dans ce champ, la plupart des substantifs attributs rencontrés se regroupent en classes de noms de comestibles, d'animaux, éventuellement d'objets (*cruche, potiche, toupie, baudruche, lavette*). Ceci se produit d'une façon tellement systématique que nous en concluons que les *acceptions familières* des noms se manifestent dans des discours qui jouent sur le croisement entre deux classes sémantiques de noms : celle qui englobe les *denotata* « naturels » d'un nom quelconque est recouverte par celle qui comporte des *designata* pour le moins inattendus, créant un effet d'incongruité.

Car, et nous touchons là un point important, ce n'est pas en tant qu'individus que les lexèmes *cruche, tarte* ou *andouille*, mis en attribution à des personnes désignées, prennent un sens familial mais *en tant qu'ils appartiennent à des classes* de dénomination sémantiquement éloignées de celles qui conviendraient pour nommer des personnes. Ce n'est qu'en vertu de cette représentativité des noms employés que les énoncés sont répétables, que les significations familières ont suffisamment de stabilité pour figurer dans les dictionnaires et servir de repères aux apprenants.

Au plan logico-sémantique, le schéma qui représente le type de nos énoncés a la forme X ÊTRE Y dans lequel X et Y symbolisent des classes sémantiques de noms. L'impertinence d'une attribution naît de la substitution d'un élément d'une classe à un élément d'une autre classe, les classes n'ayant entre elles aucun point commun. Cette formule correspond exactement à celle de la métaphore et si nos exemples livrent, effectivement, des métaphores, nous devons admettre que celles-ci sont créées par l'attribution impropre elle-même sans reposer sur la moindre analogie préexistante, contrairement à ce que pose la tradition. Nous serions bien en peine de trouver la moindre ressemblance « concrète », « objective » entre une tarte et une personne – en général – aussi stupide soit-elle ! Ceci ne dénie aucunement la circulation de multiples expressions déclassées qui sont soit motivées historiquement ou culturellement comme, par exemple : un *four* pour une représentation théâtrale sans succès, soit basées sur une ressemblance « visible » comme, toujours en visant une personne, la traiter de *dragon*, de *vache*, de *roquet*, d'*échalas* ou de *sauterelle*, etc. Le plus remarquable est que, devant le nombre quasiment illimité de possibilités d'attributions les plus diverses, imprévisibles et souvent cocasses qui s'offrent au locuteur dans ses échanges quotidiens, la langue en ait retenu autant qui ne reposent sur aucun autre fondement que sur un accord collectif toujours mystérieux. Faut-il alléguer cette « compulsion intérieure de répétition » du même qui, d'après Freud (Freud 1919 : 242 ; 1920), nous habite tous ?

En français, *traiter (de)* avec le sens de qualifier, d'appeler par tel ou tel nom, ne vaut que pour des compléments humains ou assimilés. Quant au reste, on pourra toujours parler de *navet* pour une production artistique, de *barbe* pour de l'ennui, de *pelle* pour une chute, de *pêche* pour un coup, etc., ce faisant, on requalifiera un certain référent par substitution d'un nom à un autre cependant que *traiter (de)* reste, exclusivement, un verbe de dialogue, un verbe qui implique une réponse, une réaction. Donc, traiter quiconque de *cruche*, *tarte*, *andouille* ou même de *chou* revient à le définir par les traits définitoires de son attribut – selon le schéma X ÊTRE Y comme nous l'avons précédemment souligné. Il s'ensuit qu'enfermer un être – pensant et agissant – dans les limites des caractéristiques d'un pot de terre, d'une pâtisserie, d'une fabrication charcutière ou d'un légume ne peut que révéler des sentiments malveillants ou très bienveillants de la part du locuteur et provoquer une réponse, une réaction de la part du récepteur.

Est-ce à dire que la personne désignée de la sorte dans un énoncé va conserver, désormais et à la face du monde, ses caractérisations dévalorisantes ou affectueuses ? Ce n'est, évidemment, pas le cas. Si quelqu'un est qualifié de *tarte*, ce n'est qu'en vertu d'une appréciation du locuteur, d'un jugement qui peut n'être même que provisoire, circonstanciel, le temps de la profération de l'énoncé. Cela présenterait des points communs avec l'acte d'énonciation et, de fait, nos expressions familières partagent certains traits avec les déictiques.

Cependant, nous ne sommes totalement ni dans l'anecdotique ni dans l'unicité de l'acte d'énonciation mais face à un événement énonciatif qui se répétera chaque fois que se présenteront les mêmes conditions de son apparition, conditions qui se doivent d'inclure un certain état d'esprit *définissable* du locuteur. C'est cet ensemble qui englobe les sentiments identifiés de l'énonciateur à l'égard d'autrui qui est *répétable*.

Par ailleurs, nous devons prendre en compte que nous avons, essentiellement, affaire à des énoncés-phrases nominaux. À propos de la phrase nominale, E. Benveniste (1950 : 160) écrit qu'elle exclut la subjectivité du locuteur : « Une assertion nominale, complète en soi [...] pose l'énoncé hors [...] de la subjectivité du locuteur. » Or, pour quelle raison ma conclusion aboutit-elle à l'opposé de celle de Benveniste ? Sur quel point la divergence de nos observations repose-t-elle ? L'évidence de la réponse est si flagrante qu'elle risquerait de nous échapper : tandis que Benveniste décrit le fonctionnement normé sinon *normal* du langage standard, « neutre », c'est aux emplois marginalisés par la marque *fam.* que je me suis attachée. En cela réside la différence. Là où Benveniste constate l'absence de toute subjectivité des locuteurs, je ne rencontre que des signes de sa présence. Le fait est qu'elle en vient à se manifester moyennant un total renversement des modes d'expression usuels, exploitant ces jeux que la langue autorise tout en les réprouvant, qu'elle permet en les interdisant, à peine. N'est-ce pas une des ressources proprement linguistiques que de pouvoir dire « ce qui ne se dit pas », de formuler l'interdit ?

Benveniste note encore que la phrase nominale « agit comme un argument d'autorité » (*ibid.* : 162-163), donc argument indiscutable, ajouterais-je. J'adhère pleinement à cette remarque tout en précisant que c'est toujours le locuteur qui parle, qui décrète, qui exprime *sa* vérité comme l'enfant qui croit que dire suffit pour que se matérialise son désir de puissance (Freud 1919 : 244-245).

Irais-je jusqu'à poser l'existence d'une tournure ayant valeur de déictique ? Elle présenterait alors les caractères suivants : énoncé nominal minimum, actualisé par un exclamatif, mis en fonction attributive et jouant d'un croisement entre la classe sémantique d'appartenance du *denotatum* de l'attribut et celle qui englobe les appellations courantes du référent désigné. Par ce biais s'exprimerait la subjectivité du locuteur selon ses impulsions du moment.

Bien entendu, la situation réelle intervient pour interpréter correctement l'énoncé – dans quels cas n'intervient-elle pas ? – mais, pour identifier un énoncé en tant que *familier*, point n'est besoin de recourir à l'argument pragmatique en général ; il faut et il suffit que la situation contrecarre absolument le dire, que la présence d'une quelconque cruche ou tarte ou andouille soit exclue de l'environnement de l'énonciateur lorsqu'il convoque ces vocables pour désigner l'autre, l'*alter ego*. Contrecarrer, contrarier, contrevenir, contredire, chacun de ces termes porte en lui une parcelle des possibilités d'expression de la subjectivité des êtres parlants que nous sommes. Lors de sa confrontation avec la réalité, la langue résiste en se prêtant au détournement de ses règles de fonctionnement. Mieux, en systématisant les modes de déni de ses normes, elle permet l'élaboration d'un code parallèle, lequel offre une issue aux réminiscences de désirs socialement réprimés des locuteurs.

RÉFÉRENCES BIBLIOGRAPHIQUES

- BENVENISTE Emile, 1950, « La phrase nominale », *Problèmes de linguistique générale 1*, Paris, Gallimard, 1966, p. 151-167.
- CADIOT Pierre et NEMO François, 1997, « Pour une sémiogenèse du nom », *Langue Française*, 113, p. 24-35.

- FREUD Sigmund, 1919, 1955, 1985, « L'inquiétante étrangeté », *L'Inquiétante Étrangeté et autres essais*, Paris, Gallimard, Folio-essais, p. 211-263.
- FREUD Sigmund, 1920, « Au-delà du principe de plaisir », *Essais de psychanalyse*, trad. 1920, Paris, Payot, 1968, p. 7-82.
- REY Alain, 1995, « Du discours au discours par l'usage : pour une problématique de l'exemple », *Langue Française*, 106, p. 95-120.
- SCHÖN Jackie, 1997, « Figement et régression », dans P. Fiala, P. Lafon et M.-F. Piguet (éds), *La Locution : entre lexique, syntaxe et pragmatique*, Paris, Klincksieck, p. 333-346.
- SCHÖN Jackie, 1999, « Le concept freudien d'"inquiétante étrangeté" et l'emploi "familier" des lexèmes en français », in M.-G. Pinto, J. Veloso and B. Maia (eds.), *Proceedings of the 5th Congress of the International Society of Applied Psycholinguistics (ISAPL)*, Univ. de Porto, Portugal, junho 1997, p. 207-211.
- SCHÖN Jackie, 2001, « Acceptions familières et manifestations d'affects », dans M. Arrivé et C. Normand (éds), *Linguistique et psychanalyse*, Paris, In Press, p. 157-165.
- SCHÖN Jackie, 2002, « Pour un traitement systématique des acceptions familières dans les dictionnaires », dans M. Aurnague et M. Roché (éds), *Hommage à Jacques Allières 2*, Biarritz, Atlantica-Séguier, p. 605-615.

LES MARQUES DE NIVEAUX DE LANGUE ET LA LEXICOGRAPHIE BILINGUE FRANÇAIS-GREC¹

par Efi LAMPROU (Université de Nicosie)
et Freiderikos VALETOPOULOS (Université de Poitiers)

Résumé — Dans cet article, nous nous proposons d'étudier les dictionnaires bilingues bidirectionnels français-grec. Après nous être demandé ce que l'utilisateur doit attendre d'un dictionnaire bilingue, nous donnerons un descriptif de certains dictionnaires sur support papier et examinerons leur *microstructure*, afin de savoir s'ils répondent bien aux attentes de leur destinataire. Plus précisément, nous insisterons sur l'étude de l'emploi des termes *familier*, *argotique*, *populaire* et *vulgaire*, connus sous le terme général *niveaux de langue non standard* ou *non conventionnels* et qui constituent un domaine particulièrement délicat de la lexicographie. Enfin, nous nous interrogerons sur la pertinence de certaines informations présentées dans les dictionnaires, telles que les marques assignées, ainsi que sur l'exactitude des équivalents proposés. Nous illustrerons notre démarche avec des exemples tirés de certains dictionnaires français-grec.

Mots clefs — dictionnaires bilingues bidirectionnels, niveaux de langue non standard, classes d'objet, structure argumentale, propriétés syntaxiques, classes sémantiques, combinatoire syntaxique et lexicale.

Abstract — Among the labels and features used in lexicography to reflect the heterogeneity of the lexicon, three – *Arg.* (slang), *Pop.* (popular), *Fam.* (familiar) – have attracted our attention. From both a stylistic and a sociolinguistic perspective, these labels have been much discussed, and there are scholars who believe that they can recognize and define them, and thus differentiate among them. Nevertheless, a comparative observation of the large monolingual French dictionaries clearly contradicts this; for example, the verb *jaspiner* is marked *Fam.* and *Péj.* (derogatory) in the *Petit Robert* 2003; *Pop.* in *Lexis* 2002 and *Arg.* in the *Petit Larousse* 2003. Therefore we are led to question whether these three concepts overlap to a degree according to editorial guidelines?

Keywords — bilingual dictionaries, non-formal registers, classes of objects, argument structure, syntactic properties, semantic classes, semantic and syntactic combinatory.

¹ Nous remercions Chantal Lévy pour ses conseils utiles et ses corrections.

1. LES DICTIONNAIRES BILINGUES

Le *dictionnaire bilingue de langue générale* est un dictionnaire dans lequel des unités lexicales dans une langue (dite *langue source*) sont traduites ou paraphrasées dans une autre (dite *langue cible*). Les dictionnaires bilingues sont – ou devraient être – *bidirectionnels*, c'est-à-dire aptes à servir aux deux communautés linguistiques soit de dictionnaires de version, soit de dictionnaires de thème (Marello 1997 : 34).

Les difficultés d'une telle bidirectionnalité sont plus qu'évidentes. Les premières difficultés concernent l'*anisomorphisme* des unités lexicales (voir entre autres Szende 1997). Certes, il existe certains traits définitoires qui permettent à un lexicographe ou à un traducteur de choisir telle ou telle équivalence. Par exemple, une entrée du type *τραπέζι* peut être facilement traduite par *table* dès que l'on parle de cette « pièce de mobilier composée d'une surface plane horizontale disposée à hauteur convenable pour y prendre des repas et pour y déposer les objets et les mets nécessaires à ceux-ci » (TLFI). Mais cette entrée pose quelques problèmes quand on commence à énumérer les différents mots composés ou expressions figées dans lesquels un utilisateur peut trouver cette unité lexicale : *table ronde*, *table d'honneur*, à *table*. Ainsi, l'utilisateur francophone trouvera la traduction de la *table ronde* « στρογγυλή τράπεζα » sous le mot vedette *TABLE*. L'utilisateur hellénophone aura deux solutions différentes. Tout d'abord, il sera amené à chercher sous deux mots vedettes différents, *τραπέζι* et *τράπεζα*, ce qui pose certains problèmes, car le mot *τράπεζα* apparaît dans les différents dictionnaires consultés comme polysémique, « table » et « banque ». Sinon, il devra émettre l'hypothèse que le mot *τράπεζα* est un allomorphe du mot *τραπέζι* et, par conséquent, chercher les deux mots sous le même mot vedette.

Si dans ce cas le problème vient plutôt du choix des rédacteurs, il en existe d'autres qui rendent le passage d'une langue à l'autre difficile pour des raisons sociologiques ou pragmatiques. Le cas des mots sans correspondance, c'est-à-dire des mots qui décrivent une réalité culturelle, sociale, morale inexistante dans la langue cible (*pacte civil de solidarité* ou *PACS*², *pacsé*, *ζωντοχήρα* « femme divorcée ») ou même qui décrivent une réalité qui ne correspond pas tout à fait à la réalité décrite par la langue cible (*collège* et *κολλέγιο*). À côté de ces cas relativement prévisibles, il y a aussi les unités lexicales qui ne peuvent pas être « objectivement » définies dans une autre langue. L'exemple le plus évident est celui des prédicats qui expriment une émotion. Pour commencer les mots *sentiment* et *émotion* ne correspondent pas forcément aux prédicats utilisés en grec, *αίσθημα* et *συναίσθημα*, ou même les unités lexicales qui décrivent ces émotions comme *δέος*, *φόβος*, *έρωτας* et *αγάπη* trouvent difficilement leurs équivalents en français, par exemple *crainte* et *peur* pour *δέος* et *φόβος*, *amour* pour *έρωτας* et *αγάπη*.

² En 2008, le gouvernement grec a proposé une nouvelle loi qui prévoit la création d'un *σύμφωνο συμβίωσης* « contrat de vie commune » mais uniquement pour les couples hétérosexuels.

2. LA MICROSTRUCTURE DES DICTIONNAIRES BILINGUES BIDIRECTIONNELS

La microstructure d'un dictionnaire bilingue comprend en général les informations suivantes :

- a. *L'entrée*. L'unité lexicale en position de mot vedette peut être un mot simple ou un mot composé selon la politique des rédacteurs du dictionnaire, suivi de toutes les informations grammaticales (genre, flexion et variantes graphiques) ainsi que de sa transcription phonétique.
- b. *Les acceptions et les traductions dans la langue cible*. Chaque acception est accompagnée des informations suivantes : toutes les marques nécessaires à préciser le niveau de langue ou le domaine d'emploi, les propriétés syntaxiques et, enfin, la traduction ou la paraphrase suivie d'un exemple suffisamment clair pour que l'utilisateur puisse percevoir les propriétés syntaxiques de l'entrée. Parfois, le dictionnaire présente la combinatoire lexicale de l'entrée, comme par exemple les verbes, les noms ou les adjectifs appropriés³. Par exemple, le français utilise avec *table* au sens de « la table à l'occasion d'un repas ou pendant un repas » (TLFI), les verbes *quitter*, *sortir de* et *se lever de*, contrairement au grec qui utilise uniquement *σηκώνομαι από το τραπέζι* « se lever de table ».

Il est évident que les informations choisies peuvent différer selon le type et les objectifs du dictionnaire. Ainsi, un dictionnaire de poche peut limiter ses informations d'ordre syntaxique ou ses exemples.

Examinons maintenant la microstructure de la partie français-grec de Lust et Pantelodimos (1995), qui présente la structure de ses entrées en introduction (voir en page suivante – ce schéma ne figure plus dans l'édition 2004). Dans ce dictionnaire, les rédacteurs ont décidé de donner toutes les informations vues ci-dessus : la transcription phonétique, les variantes graphiques, la catégorie grammaticale et des précisions supplémentaires, le champ sémantique, les différents sens, les marques pour indiquer le niveau de langue ou s'il s'agit d'une expression libre ou figée.

Contrairement à sa partie français-grec, le dictionnaire grec-français de la même maison d'édition ne respecte pas cette structure. Ainsi, la transcription phonétique du grec n'apparaît pas et les informations concernant la catégorie grammaticale ne sont pas toujours mentionnées, sauf pour les verbes et parfois pour le participe passé qui peut apparaître comme mot vedette avec un renvoi vers l'article du verbe. Par exemple :

αγυροξυπνημένος⁴, -η, -ο *μτχ.* ☞ αγυροξυπνώ

Le deuxième dictionnaire de notre étude, celui de Rosgovas (2002), possède la même microstructure dans les deux directions du dictionnaire. Dans ce dictionnaire, le rédacteur présente la transcription phonétique du grec et du français, les variantes graphiques, la catégorie grammaticale et des précisions supplémentaires, le champ sémantique, les différents sens séparés à l'aide d'un signe typographique, les marques pour indiquer le niveau de langue ou s'il s'agit d'une expression libre ou figée. Les différences de sens sont indiquées à l'aide de la traduction mais il n'y a pas d'informations sur les propriétés syntaxiques.

3 Un prédicat est approprié quand il est spécifique d'une classe d'arguments. Ainsi, le verbe *rédiger* est le verbe approprié de la classe <texte> (Le Pesant & Mathieu-Colas 1998 : 12).

4 Quelqu'un qui vient d'être réveillé, qui est réveillé de bonne heure ou qui est mal réveillé.

MODE D'EMPLOI DU DICTIONNAIRE ΔΙΑΓΡΑΜΜΑ ΧΡΗΣΗΣ ΤΟΥ ΛΕΞΙΚΟΥ



3. LES MARQUES DE NIVEAUX DE LANGUE ET LES DICTIONNAIRES BILINGUES

La définition de *registres de langue* a fait l'objet d'une longue discussion ; de nombreux linguistes et littéraires ont centré leurs études sur les différents styles – un terme assez vague – ou modes d'écrire ou de parler. Si l'on veut donner une définition très générale, un *registre de langue* est un style d'expression adapté à une situation de communication particulière. Ce style détermine certains choix linguistiques qui concernent le lexique et la syntaxe ou même la prononciation. Plusieurs classifications ont été proposées qui distinguent habituellement les trois registres suivants : *soutenu*, *courant* et *familier* (voir entre autres Gadet 2003 : 99). Nous rencontrons également le registre *vulgaire* et le registre *populaire*. Parmi ces niveaux de langue, nous nous concentrerons sur les niveaux dits *non standard* ou *non conventionnels*, c'est-à-dire les niveaux *familier*, *vulgaire*, *populaire* et *argotique*.

Dans la suite de notre article, nous présenterons tout d'abord une brève définition des marques que nous avons étudiées. Ensuite, nous présenterons l'utilisation de ces marques dans certains dictionnaires bilingues en comparaison avec deux dictionnaires monolingues, français et grec, qui nous ont servi de référence.

3.1 LES MARQUES DE NIVEAUX DE LANGUE NON CONVENTIONNELS

L'*argot*⁵ est un lexique particulier utilisé par un groupe socioprofessionnel généralement homogène et plus ou moins restreint. Le locuteur affiche son appartenance au groupe en question et se distingue de la grande majorité des sujets parlants. Denise François-Geiger souligne les fonctions cryptique, ludique ou même ludico-cryptique de l'argot et mentionne ses autres rôles qui relèvent du snobisme ou même de la pudeur (François-Geiger 1991, voir aussi les travaux sur les argots et les « français populaires » dans *Marges Linguistiques*, 6, 2003). Pour présenter quelques exemples provenant du grec moderne, nous pouvons recourir à l'argot militaire⁶ et mentionner les mots suivants :

- *ροζαλία*, le certificat du service militaire ; le nom est formé sur la base *ροζ* « la couleur rouge pâle » à cause de la couleur du document.
- *θαλαμοντόγκ* ou *θαλαμόσκυλο* « mot composé des mots *dortoir* + *chien* ». Le mot signifie celui qui est chargé de la garde du dortoir.
- *ψάρι* ou *ψαρούκλα* « poisson ou gros poisson », celui qui vient de commencer son service militaire.

5 La raison pour laquelle nous incluons le vocabulaire argotique dans cette liste est le fait que cette marque est utilisée par les dictionnaires bilingues pour décrire certaines unités lexicales qui ont la marque *vulgaire* ou *populaire* dans le dictionnaire monolingue. Par exemple, dans le *Dictionnaire grec moderne-français* de D. Pantelodimos et K. Kaïteris, l'entrée *αβέρτα* est marquée *αργκ.* (argotique) pour l'emploi *μιάω αβέρτα* « parler trop » : *αβέρτα επιρρ.* (λαϊκ.) sans retenue, sans réserve, sans compter. 1. *Ξοδεύω αβέρτα λεφτά*. Flamber de l'argent. 2. (αργκ.) *Μιλάει αβέρτα, θα το φάει το κεφάλι του !* Il parle trop, il va (finir par) s'en mordre les doigts !

6 Pour l'argot militaire, voir les sites pour les jeunes comme <http://www.skopies.net/lexi/index.html> (dernière visite 01.08.2007). Voir aussi tous les travaux d'Elias Petropoulos, et plus précisément l'argot des homosexuels ou l'argot des jeunes bourgeois, un inventaire qu'il a terminé juste avant sa mort en 2003. Plusieurs sites sur l'argot et la langue des jeunes ont vu le jour ces cinq dernières années.

- *γκοτζίλας* « adaptation en grec du nom Godzilla », qui renvoie à la viande servie à l'armée. Le nom fait référence à la mauvaise qualité de la restauration.

Le registre *familier*, un terme qui n'est pas sans poser de problèmes, concerne le vocabulaire, la prononciation et la grammaire. Le locuteur utilise un terme familier inconsciemment et s'adapte à la situation de communication (Szabó 2001). À titre d'exemple, nous pouvons mentionner *μασαμπουκώνω* ou *περιδρομιάζω* pour *bouffer*. Les mots utilisés ne semblent pas connoter un groupe social quelconque.

Le registre dit *populaire* est encore plus problématique que le registre *familier*. Il semble que l'emploi d'un terme populaire dépend plutôt de l'origine sociale et de l'éducation du locuteur (Szabó 2001). Par exemple, la forme *κόκκαλο* « bouche bée » est plus marquée que la forme *εμβρόντητος* « sidéré, ébahi ».

Enfin, la notion de *vulgaire* est très ambiguë et renvoie aux *gros mots*, aux *obscénités*, aux *mots tabous*, c'est-à-dire à tous les emplois de nature vulgaire, voire « choquante ». L'assignation de cette marque implique un jugement subjectif : la marque peut caractériser soit le signifié soit le signifiant. Ainsi, dans le dictionnaire monolingue grec de Manolis Triantafylidis, l'unité lexicale *σκατίλα* « l'odeur de merde, ça sent la merde » est marquée comme *vulgaire* contrairement à la forme *ποδαρίλα* « l'odeur de pieds, ça sent les pieds ». Par contre, entre les deux synonymes *πέος* « pénis » et *πούτσα* « pénis, queue », le premier n'est pas marqué contrairement au second qui a la marque *vulgaire* à cause de sa forme et non plus à cause de son signifié.

Par conséquent, nous devons souligner que l'attribution d'une définition aux registres peut être relativement simple, mais, à l'usage, l'application de ces registres est une affaire beaucoup plus complexe. Les raisons en sont multiples : l'intégration progressive des unités argotiques dans la langue générale, l'évolution des mœurs qui atténue la force des mots anciennement rejetés, etc.

3.2 LES DICTIONNAIRES CONSULTÉS

3.2.1 LES DICTIONNAIRES BILINGUES FRANÇAIS-GREC – GREC-FRANÇAIS

Dans la suite de notre article, nous présenterons les dictionnaires bilingues que nous avons choisis pour notre étude. Le critère qui a guidé notre choix était tout d'abord la bidirectionnalité des dictionnaires. Les dictionnaires choisis devaient comporter les deux parties français-grec et grec-français. Ce critère a écarté de notre étude le dictionnaire de J.-P. Robert et M. Malamas-Robert (2002), *Dictionnaire français-grec moderne*, qui n'a pas, à notre connaissance, de deuxième partie, ou le *Σύγχρονο Ελληνογαλλικό Λεξικό* de Galanis (1998). Le deuxième critère concernait la langue grecque décrite dans le dictionnaire. Nous avons fait le choix de présenter les dictionnaires qui décrivent la langue grecque contemporaine. Pour cette raison, nous avons exclu de notre étude le dictionnaire *Τέλειο Ελληνογαλλικό - Γαλλοελληνικό Λεξικό*, dit Mandeson (1972) qui utilise la *katharevousa*⁷ ainsi que le système polytonique, c'est-à-dire les diacritiques abandonnés depuis plus d'une trentaine d'années.

⁷ Il s'agit d'un idiome artificiel et puriste imitant une représentation idéalisée du grec attique.

3.2.2 LES DICTIONNAIRES MONOLINGUES FRANÇAIS ET GREC

Avant de passer aux dictionnaires bilingues et aux problèmes rencontrés, il serait utile de présenter les dictionnaires monolingues grecs et français que l'on a employés pour comparer les marques lexicographiques rencontrées.

Nous avons sélectionné le dictionnaire de Manolis Triantafilidis pour des raisons aussi bien pratiques que théoriques. La version électronique⁸ du dictionnaire nous a permis de faire des recherches approfondies et de sélectionner les entrées selon leur marque. Par ailleurs, les rédacteurs du dictionnaire papier présentent dans l'introduction les marques choisies ainsi qu'une courte description pour expliquer aux utilisateurs l'emploi de chacune (voir aussi l'article d'A. Anastassiadis-Syméonidis, p. 53 et suiv. dans ce volume). Nous avons donc relevé les marques qui nous intéressent et leur définition :

- *οικείο* (familier) : se dit des mots utilisés en famille et dans des situations de communication amicales. Le locuteur évite d'employer ces mots quand il s'adresse à un inconnu ou à un supérieur dans une hiérarchie professionnelle ou autre. Par exemple, *το καθισιό* « le farniente », *ο παλιόφιλος* « un ami de longue date ».
- *λαϊκό* (argotique) : se dit des mots souvent de la langue orale, utilisés par des gens vivant en marge de la société ou par des jeunes. Ces mots peuvent être utilisés également dans la langue courante. Par exemple, *το πασαπόρτι* « passeport », mais utilisé dans des expressions comme *δίνω το πασαπόρτι* « chasser qqn », *η παρόλα* « discours pompeux ».
- *χυδαίο* (vulgaire) : se dit des gros mots, des mots tabous, liés à la sexualité ou aux besoins physiques. Par exemple *η κουράδα* « la crotte ».
- *λαϊκότροπο* (populaire) : se dit des mots dialectaux qui peuvent être connus par les habitants des grandes agglomérations et qui sont utilisés dans la littérature. Par exemple, *το δίστρατο* « la fourche ».

Le *Nouveau Petit Robert* (2007 : xv-xvj) traite des différentes marques sous l'appellation de « niveaux de langue » qui, selon cette description, font partie de *variétés du français*. Les auteurs soulignent qu'« elles sont signalées avec le plus de précision possible, compte tenu de l'ampleur de la description et de l'absence de tabous » (2007 : xvj). En considérant les niveaux de langue comme un continuum, ils distinguent les emplois réservés à la langue écrite et à des discours « soutenus » caractérisant le français écrit (*littér.*) des emplois « familiers » (*fam.*) tout en soulignant qu'il y a, entre ces deux axes, diverses nuances possibles. Dans ce sens, les usages propres à une classe sociale donnée (*arg.* pour *argot* ou *argotique*) sont ou doivent être bien distingués des emplois qui sont liés à une scolarisation insuffisante dans certains milieux sociaux défavorisés (*pop.* pour *populaire*). De la même manière, la marque *vulgaire* (*vulg.*), qui, selon la même description, est utilisée dans des contextes spéciaux et pour des mots qui sont susceptibles de choquer, se différencie du *péjoratif* (*péj.*), qui caractérise une attitude violente comme par exemple une insulte ou une injure⁹.

8 <http://www.komvos.edu.gr/dictionaries/dictonline/DictOnLineTri.htm>

9 Les termes d'injure raciste ou sexiste sont d'autant plus signalés que leur emploi public constitue, en France, un délit susceptible de poursuites judiciaires devant les tribunaux correctionnels.

3.3 LES MARQUES DANS LES DICTIONNAIRES FRANÇAIS-GREC

3.3.1 LE DICTIONNAIRE FRANÇAIS-GREC MODERNE

DE C. LUST ET D. PANTELODIMOS (2004)

Le dictionnaire de Lust & Pantelodimos (2004) contient 140 000 entrées et 1 112 pages. Les auteurs semblent soucieux des différents niveaux de langue et des problèmes de traduction. Dans leur préface¹⁰, ils soulignent que les dictionnaires les plus anciens doivent être utilisés avec prudence : certains sont complètement dépassés, d'autres se limitent à donner « un simple équivalent grec des mots français sans proposer ni exemples, ni expressions idiomatiques, et sans non plus indiquer le niveau de langue ». Concernant notre sujet, nous constatons un peu plus loin que pour la rédaction de ce dictionnaire ont été retenus « le vocabulaire de base des discours écrit et oral contemporains, la langue familière quotidienne ainsi que les expressions idiomatiques courantes. Par ailleurs, les mots d'argot, voire, mais plus rarement, les expressions triviales, n'ont pas été censurés, dès lors que leur emploi s'est généralisé ». Ainsi, dans la liste des signes et des abréviations, nous retrouvons *arg.* (pour argot ou argotique), *fam.* (pour familier), *vulg.* (pour vulgaire) et enfin, *pop.* (pour populaire). Il faut souligner que les rédacteurs ne définissent pas les marques utilisées.

Malgré cette introduction prometteuse, l'étude de la macrostructure et de la microstructure du dictionnaire révèle certains problèmes et incohérences par rapport à tout ce qui a été annoncé. Un des premiers problèmes que nous constatons concerne l'emploi des marques utilisées dans ce dictionnaire et les marques utilisées dans les dictionnaires monolingues. Prenons l'exemple du mot *amour*. Dans le *Petit Robert* (2007), la locution *faire l'amour* n'est pas marquée, comme par ailleurs dans le *TLFI*, contrairement à *être un amour* qui est marqué *familier par extension*. Par contre, dans le dictionnaire bilingue en question l'expression *faire l'amour* est marquée comme populaire.

AMOUR n. m. [...] || 2. *pop.* *Faire l'amour.* Κάνω έρωτα. [...] || *fam.* *Tu serais un amour de m'apporter mes lunettes.* Θα είσαι γλύκα, αν μου φέρεις τα γυαλιά μου.

Il faut noter que l'entrée équivalente dans le dictionnaire grec-français reprend la même traduction pour l'expression *κάνω έρωτα* sans marque, ce qui est également le cas du dictionnaire monolingue grec.

Un deuxième problème concerne la cohérence dans l'emploi des marques par rapport à la traduction proposée. Ainsi, l'entrée marquée *fam.* en français est traduite par une unité lexicale qui n'est pas marquée en grec. Pourtant, dans la microstructure, les auteurs recourent à des exemples qui contiennent des unités lexicales marquées en grec. Par exemple :

PIGER v. tr., (*fam.*) καταλαβαίνω. 1. *Il pige tout de suite.* Αμέσως μπαίνει στο νόημα. 2. *Je ne pige rien à la physique.* Δεν σκαμπάζω τίποτα, γρι, από Φυσική. 3. *Tu as pigé ?* Μπήκες?

σκαμπάζω [skambázo] P2.1α : (προφ.) έχω κάποιες γνώσεις πάνω σε κτ., καταλαβαίνω : [...] (*Manolis Triantafyllidis*)

10 Nous devons souligner que les préfaces des deux éditions présentent des modifications significatives.

Ainsi, les rédacteurs choisissent-ils le verbe *καταλαβαίνω* pour traduire le verbe *piger*, alors que dans les exemples utilisés ils recourent au verbe *σκαμπάζω* marqué *προφορικό* (familier) dans le dictionnaire monolingue grec.

3.3.2 LE DICTIONNAIRE FRANÇAIS-GREC DE ROSGOVAS (2002)

L'auteur de ce dictionnaire des éditions Rosgovas ne mentionne pas le nombre d'entrées, soulignant que celui-ci « n'apporte rien si la présentation, la fonctionnalité, le vocabulaire de base et surtout les expressions lui font défaut ». Dans l'introduction, il n'y a aucun renseignement sur les marques utilisées ; un index placé au début du dictionnaire ne présente que les deux marques *λαϊκά* « populaire » et *κοινά*¹¹, qui semblent avoir la fonction d'une marque générique.

piger ρ.μ. (λαϊκ.) καταλαβαίνω. / σκαμπάζω // t'as pigé ? = καταλαβού? / μπήκες.

gosse v. (κοιν.) πιτσιρίκι // beau – = ωραίο παιδί / belle – = ωραία κοπέλα // sale – = παλιόπαιδο.

Mais nous constatons assez vite que cette marque n'est pas régulièrement utilisée. Ainsi, l'entrée *gamin* n'est pas marquée contrairement aux formes *môme*, qui est marquée *λαϊκό*, et *gosse*, qui porte la marque *κοινό*. Le *Petit Robert* (2007) accorde la marque *fam.* aux entrées *gamin* et *gosse* ainsi qu'à l'entrée *môme* dans le sens *enfant*.

3.4 LES DICTIONNAIRES GREC-FRANÇAIS

3.4.1 LE DICTIONNAIRE GREC MODERNE-FRANÇAIS DE D. PANTELODIMOS ET K. KAITERIS (2002).

Le dictionnaire de Pantelodimos et Kaiteris, 150 000 entrées et 1 507 pages, suit l'exemple de la version français-grec du dictionnaire qui « a comblé, de l'avis général, un grand vide dans la bibliographie lexicographique » (2002 : ix). D'après les auteurs,

le dictionnaire [...] a également l'ambition de mettre en valeur le trésor des deux langues et de démontrer la richesse lexicale, phrastique et sémantique, du français et de la langue néo-hellénique dans leur état actuel. C'est pourquoi il comprend pratiquement tout l'ensemble du vocabulaire grec utilisé pour la communication langagière, orale ou écrite, à tous les registres : littéraire, officiel, familier, quotidien, sans en exclure le code sentimental de l'argot pour des mots dont l'utilisation est aujourd'hui généralisée.

Par ailleurs, ils notent un peu plus loin qu'« un effort substantiel a également été fourni pour harmoniser les niveaux de langue, afin que soient conservées les équivalences de registre [...] ». Dans le cas où la coïncidence des proverbes n'était pas possible, ils ont « jugé nécessaire de rendre les idiotismes et les expressions courantes de sorte que leur sens, esprit et nuance restent sensibles dans l'autre langue ». Dans le cas contraire, les auteurs ont donné une définition ou une description explicative.

L'étude du dictionnaire a révélé certains problèmes. Plusieurs mots marqués dans le dictionnaire bilingue comme *populaire* ou autre n'ont pas vraiment de marque dans les dictionnaires monolingues. Par exemple, le mot *αβαρία*

¹¹ Les exemples ne permettent pas de comprendre le sens de *κοινά*. Nous supposons qu'il s'agit de *familier*.

« avarie » est marqué *λαϊκ.* « populaire » pour des raisons qui ne sont ni claires ni évidentes :

αβαρία, η (λαϊκ.) avarie *f*, dommage *m*, dégât *m*, compromis *m*, concession *f*.
1. [...]

Dans les dictionnaires monolingues, le même mot ne porte aucune marque :

αβαρία η [avaría] O25 : 1. (ναυτ.) α. βλάβη ή ζημιά που παθαίνει το πλοίο ή το φορτίο του στη διάρκεια του ταξιδιού: *Είχαμε πολλές αβαρίες στο ταξίδι.* β. [...] (*Manolis Triantafylidis*)

Nous pouvons aussi avoir le cas inverse, c'est-à-dire que l'entrée n'a aucune marque dans le dictionnaire bilingue contrairement à celle du dictionnaire monolingue. Par exemple, le mot *παράς* est marqué comme *προφ.* (fam. / oral) dans son acception 2. *argent*, ce qui n'est pas noté dans le dictionnaire bilingue :

παράς, ο [...] 2. *Έχω παράδες.* Avoir de la galette. *Έχω παρά με ουρά.* Avoir de l'argent à la pelle...

παράς ο [parás] O1 : 1. ... 2. (προφ.) το χρήμα, τα λεφτά: *Έχει / βγάζει / κερδίζει / παίρνει (πολλούς) παράδες.* [...] (*Manolis Triantafylidis*)

Par ailleurs, nous remarquons que l'attribution de marques n'est pas régulière. Par exemple, le verbe *παίρνω* « prendre » apparaît dans un certain nombre de locutions sans aucune marque : *παίρνω αέρα* « monter à la tête », *παίρνω κάποιον στο ψιλό* « entuber quelqu'un », *πάρ'τα* « prends ça pour toi (accompagné d'un geste obscène) » ou *παίρνω ανάποδες* « mon sang ne fait qu'un tour ». Contrairement aux expressions ci-dessus, les seules qui soient marquées sont *τα παίρνω* « se faire graisser la patte », qui est marquée *familier*, et *τον παίρνει* « elle fait l'amour, elle se fait mettre », marquée *péjoratif*.

Enfin, l'attribution de telle ou telle autre marque peut être aussi arbitraire et sans cause évidente. L'exemple que nous avons déjà mentionné (*supra*, n. 6) est celui de l'entrée *αβέρτα* qui est marqué de *αργκ.* « argotique » dans l'acception 2. *μιλάει αβέρτα, θα το φάει το κεφάλι του !* « Il parle trop, il va (finir par) s'en mordre les doigts ! ». Il s'agit d'une phrase qui pourrait être utilisée dans plusieurs contextes différents sans désigner pour autant un vocabulaire particulier employé au sein d'un certain groupe socioprofessionnel (voir aussi la définition, section 3).

Il en est de même pour toute la liste des mots formés sur la base nominale *πούστης* « pédé », marqués comme *υβριστικό* (*péjoratif*). Les noms qui désignent un humain sont marqués comme *υβριστικό* (*péjoratif*) contrairement aux autres unités lexicales, par exemple *πουστιά* « crapulière, saloperie », *πούστικα* « de façon dégueulasse », *πούστικός* « de pédé ». Ces exemples nous permettent de souligner le problème d'attribution d'une marque qui se fait souvent à l'aide de critères fondés sur une certaine représentation du monde extralinguistique.

3.4.2 LE DICTIONNAIRE GREC-FRANÇAIS DE ROSGOVAS (2002)

Contrairement à la partie français-grec, la partie grec-français utilise les abréviations *fam.* pour familier et *pop.* pour populaire. L'auteur n'explique pas dans son introduction les raisons pour lesquelles il utilise les deux marques et l'utilisateur ne sait pas vraiment à quoi chacune correspond. Quelques exemples relevés dans le dictionnaire ne nous permettent pas d'en comprendre plus.

Ainsi la traduction « chatouille » pour le nom *γαργαλητό* est marqué *pop.* contrairement au nom *γαργάλημα* « chatouillement ». Le dictionnaire monolingue de Manolis Triantafylidis n'accorde aucune marque à aucun des deux et le TLFi marque de *fam.* l'expression *faire des chatouilles*. Pour finir, il faut noter que le terme *familier* est extrêmement rare dans ce dictionnaire.

4. LES MARQUES DE NIVEAUX DE LANGUE : QUELLE POLITIQUE ?

Après cette présentation qui nous a permis de souligner quelques problèmes de la lexicographie bilingue, nous avons donc envisagé l'approche proposée par la théorie des *classes d'objets* (Langages 131) qui a pour objectif la description de la langue en vue de la traduction automatique.

Nous étudierons le cas des prédicats qui dénotent un conflit verbal comme *dispute*, *chicane*, *prise de bec*, *brouille*, *engueulade*, *chamaillerie*, etc. qui sont présentés comme des synonymes dans les dictionnaires de langue mais qui ont un comportement syntaxique différent selon leur niveau de langue. Bien qu'ils ne se distinguent pas par l'actualisation – ils sélectionnent, en général, tous le verbe support *έχω* « avoir »¹² – les prédicats *engueulade*, *chamaillerie*, *bisbille*, *prise de bec*, caractérisés par la marque *fam.*, se différencient de *dispute*, *chicane* ou *discorde* par deux propriétés syntaxiques : la nature de leurs arguments¹³ et les adjectifs appropriés.

En ce qui concerne leurs arguments, *καυγαδάκι*, *καυγός*, *τσακωμός*, *κατσάδα* / *engueulade*, *chamaillerie*, *bisbille*, *prise de bec* apparaissent moins fréquemment dans des phrases comme :

- (1) ??Ο έλληνας πρωθυπουργός είχε ένα καυγαδάκι με το γάλλο ομόλόγο του σχετικά με το θέμα της ονομασίας των Σκοπίων
Le grec premier ministre a eu une engueulade avec le français homologue son en ce qui concerne le sujet du nom de Scopie¹⁴
- (2) ??Ο έλληνας πρωθυπουργός τα τσούγκρισε με το γάλλο ομόλόγο του σχετικά με το θέμα της ονομασίας των Σκοπίων
Le grec premier ministre a été en bisbille avec le français homologue son en ce qui concerne le sujet du nom de Scopie

ce qui n'est pas le cas pour les prédicats *διαπληκτισμός*, *φιλονικία*, *διένεξη* / *dispute*, *discorde* ou *chicane* :

- (3) Ο έλληνας πρωθυπουργός είχε (έναν διαπληκτισμό + μια φιλονικία + μια διένεξη) με το γάλλο ομόλόγο του σχετικά με το θέμα της ονομασίας των Σκοπίων

12 D'autres verbes supports plus spécifiques, appropriés à tel ou tel prédicat, sont également valables. C'est par exemple le cas du verbe support *être* dans la locution *être en bisbille avec qqn* qui a comme équivalent le grec *τα έχω τσούγκρισει με κάποιον* « se chamailler avec qqn », de niveau de langue *fam.* selon certains des dictionnaires consultés. Comme il a été déjà dit au § 1, nous constatons une certaine incohérence entre les dictionnaires bilingues utilisés pour cette étude. Ainsi, pour l'entrée *bisbille*, la marque *fam.* apparaît dans le dictionnaire français-grec, alors que cette dernière disparaît lorsque nous consultons la même entrée dans le dictionnaire grec-français.

13 Il s'agit de traits syntactico-sémantiques tels que *humain*, *non humain*, *animal*, *locatif*, *végétal*, *inanités concrets*, les noms de *temps*. Pour une analyse plus détaillée, voir *Langages*, 131.

14 Nos exemples sont suivis d'une traduction littérale en français.

Le grec premier ministre a eu une (dispute + discorde + chicane) avec le français homologue son en ce qui concerne le sujet du nom de Scpie.

Les conflits verbaux *dispute*, *discorde* ou *chicane*, liés à des mésententes entre des personnes qui n'entretiennent pas des rapports familiaux et affectifs, et qui ne surviennent généralement pas pour un motif futile, peuvent sélectionner comme arguments des noms tels que *ministre*, *premier ministre*, *directeur*, des noms de pays, de représentants de partis, etc.

Quant aux adjectifs appropriés qui caractérisent cette classe, si le prédicat *dispute* se combine parfaitement avec l'adjectif *parlementaire* (4), ce dernier s'associe difficilement avec *καβγαδάκι*, *τσακωμός* / *engueulade*, *chamaillerie*, etc. (5) :

- (4) *Η κοινοβουλευτική (διένεξη + φιλονικία)*
La parlementaire (dispute + discorde)
- (5) **Ο κοινοβουλευτικός τσακωμός) + (*το κοινοβουλευτικό καβγαδάκι)*
La parlementaire (engueulade + chamaillerie).

Les verbes appropriés constituent une autre propriété syntaxique permettant de distinguer les prédicats qui n'ont pas le même niveau de langue, et qui sont utilisés dans des situations communicationnelles différentes. Nous avons appliqué ce critère non pas aux <disputes>, mais aux prédicats nominaux dénotant un autre type d'échanges à caractère verbal, à savoir les <discussions>. Les verbes *animer* ou *organiser* / *διευθύνω*, *οργανώνω* se combinent avec les noms *débat*, *dialogue*, *discussion*, *entretien*, *συνομιλία*, *συνδιάλεξη*, *ομιλία*, *τηλεοπτικός διάλογος* dans les deux langues, ce qui n'est pas le cas pour les échanges verbaux tels que *κουβέντα*, *κουβεντολόι*, *κουβεντούλα*, *ψιλοκουβεντούλα*¹⁵, *πάρλα* / *petite conversation*, *parlotte*, *causette*, etc. (6 et 7). Les premiers décrivent des discussions organisées, les seconds des discussions informelles. Selon les dictionnaires consultés, ces derniers portent la marque *fam.* dans les deux langues :

- (6) *France Info a organisé (un débat + un dialogue + une discussion + *une causette + *une parlotte + *une bavette) sur les produits OGM*
- (7) *Γνωστός δημοσιογράφος διήθνε τη (συζήτηση + τηλεοπτική αναμέτρηση + *κουβέντα + *ψιλοκουβέντα) σχετικά με το ασφαλιστικό*
Réputé journaliste a animé la (discussion + débat télévisé + *causette + *parlotte) concernant la sécurité sociale.

Inversement, les verbes tels que *faire* ou *tailler* ne peuvent se combiner qu'avec les prédicats qui décrivent une discussion à caractère informel (8) en français, alors qu'ils sont complètement inappropriés pour *débat*, *dialogue*, *discussion*, etc. (9) :

- (8) Il a fait (la parlotte + la causette¹⁶ + *le débat + *la discussion) avec la voisine
- (9) Il a taillé (une bavette + *un débat + *un dialogue diplomatique) avec sa voisine.

15 *Faire un brin de causette* est l'expression équivalente en français. Le grec dispose d'une forme nominale, *ψιλοκουβέντα*, ce qui n'est pas le cas du français.

16 *Faire la parlotte* et *faire la causette* sont marqués *fam.* dans le *Petit Robert* (2008).

De même, en grec, les verbes qui se combinent avec les noms *κουβέντα*, *κουβεντολόι*, *κουβεντούλα*, *ψιλοκουβεντούλα*, *πάρλα*, etc. / *petite conversation*, *parlotte*, *causette*, *tchatche* ne peuvent pas s'associer avec les prédicats nominaux de discussions à caractère officiel tels que *συνδιάλεξη*, *ομιλία*, *τηλεοπτικός διάλογος* / *conversation*, *entretien*, *débat*, *débat télévisé* (10 et 11). Il s'agit en effet d'échanges verbaux quotidiens. Cette incompatibilité est confirmée par la nature différente de leurs arguments :

- (10) *Επιασε πάλι (κουβέντα + κουβεντούλα + το κουβεντολόι + *συνδιάλεξη + *ομιλία) με τις γειτόνισσες*
A taillé encore la (causette + parlotte + *conversation + *entretien) avec les voisines
- (11) *Αρχισε την πάρλα και δεν έλεγε να σταματήσει*
A commencé la parlotte et il n'était pas question qu'il arrête.

La nature et le schéma des arguments¹⁷ arrivent non seulement à distinguer les prédicats nominaux de parole <discussions> entre eux, mais aussi les verbes de cette classe. De même, les « petites conversations », le plus souvent à caractère péjoratif, possèdent une syntaxe différente.

Dans un premier temps, nous avons comparé les verbes tels que *συνδιαλέγομαι*, *συνομιλώ*, *διαπραγματεύομαι* / *dialoguer*, *s'entretenir*, *débattre*, etc. avec les verbes du type *τα λέω με κπ*, *λέω δυο λόγια* / *tchatcher*, *causer*, etc. Ils traduisent tous l'action de discuter. Pourtant, ils ne peuvent pas être utilisés dans la même situation communicationnelle, ce qui se confirme aussi par la syntaxe. Premièrement, *dialoguer*, *s'entretenir*, *débattre*, etc., qui traduisent le fait de s'entretenir avec quelqu'un sur un sujet particulier et d'une manière organisée, sélectionnent comme arguments des humains tels que *chef d'État*, ce qui semble moins évident pour les verbes *tchatcher* ou *causer* et leur équivalent grec :

- (12) Les deux chefs d'État (se sont entretenus de + ?*ont tchatché sur + ?*ont causé de la crise économique mondiale).

Deuxièmement, le troisième argument de la structure argumentale, c'est-à-dire « l'objet » de la conversation, semble plus obligatoire pour les discussions organisées que pour les échanges verbaux *tchatcher* et *causer* qui dénotent l'action de s'entretenir familièrement avec quelqu'un, parler de choses et d'autres :

- (13) Les deux chefs d'État (se sont entretenus de + ?*ont tchatché sur¹⁸ + ?*ont causé de la crise économique mondiale).

Ces derniers, contrairement aux verbes *dialoguer*, *s'entretenir*, *débattre*, etc., semblent, en plus, sélectionner difficilement comme arguments les circonstan-

17 Selon les travaux de Z. S. Harris (1976), l'unité minimale de sens est la phrase simple constituée par un prédicat et ses arguments. Le prédicat (verbe, nom, adjectif ou préposition) sélectionne ses arguments (le sujet et ses compléments). Pour décrire exhaustivement et systématiquement le lexique dans la perspective du traitement automatique de langues, Gaston Gross (1998) introduit, en plus, la notion d'*emploi*, selon laquelle les prédicats sont décrits en fonction de toute une série de propriétés syntaxiques (entre autres, le schéma d'arguments, i.e. *préd (arg1, arg2, arg3)*, l'actualisation, les informations aspectuelles, le niveau de langue, etc.). Le schéma d'arguments du verbe *manger* par exemple sera le suivant : *manger* (<hum>, <aliment>).

18 Selon le PR, le verbe *tchatcher* est intransitif. Pourtant, on trouve sur google.fr un grand nombre d'occurrences dans lesquelles *tchatcher* est transitif.

ciels qui traduisent le *lieu*, le *temps* et le *témoin*¹⁹ (14a et b). Ce type de phrases semble peu naturel. La raison en est que les prédicats qui décrivent l'action de parler de choses et d'autres ne peuvent pas, comme les discussions organisées, avoir un emploi événementiel ; c'est pourquoi ils ne partagent pas la syntaxe des *événements* :

- (14a) *Les deux chefs d'État (ont tchatché + ont causé) sur la crise économique mardi soir à Bruxelles en présence de leurs ministres et conseillers
 (14b) Les deux chefs d'État se sont entretenus mardi soir à Bruxelles de la crise économique mondiale en présence de leurs ministres et conseillers.

On constatera que pour le deuxième type d'échanges verbaux, les prédicats verbaux qui se rapprochent des discussions tout en s'en différenciant dans leur majorité par leur caractère péjoratif, se distinguent aussi entre eux selon leur niveau de langue (15-17). Les verbes *baver*²⁰, *bavasser*, *débiner*, *taper*²¹, *causer*²¹, *rapoter*, etc. ou bien *κουτσομπολεύω*, *ανοίγω πηγαδάκι*, *βγάζω αφρούς*, *θάβω* κπ., *κακολογώ* κπ., *ρίχνω λάσπη* σε κπ., etc. ne peuvent pas se trouver dans le même contexte que les verbes *bavarder*, *dénigrer*, *critiquer*, *médire*, *φλυαρώ*, *επικρίνω*, *δυσφημώ*, ce qui se traduit syntaxiquement par le fait qu'ils ne sélectionnent pas le même type d'arguments :

- (15) (*Ο Πέτρος* + **ο διευθυντής*) *θάβει συνεχώς το διπλανό του*
 (Le Petros + directeur) débine constamment son voisin
 (16) *Οι συμμαθητές του τον (κουτσομπολεύουν + κακολογούν)*
 Les camarades ses le (cancanent + bavassent) taper sur qqn (PR, fam.)
 (17) *Οι συνάδελφοι του τον (δυσφημούν + επικρίνουν)*
 Les collègues ses le (dénigrent + médisent).

L'emploi des noms et des verbes précités s'avère alors assez problématique pour les apprenants et pour les utilisateurs hellénophones de dictionnaires bilingues. En mettant ce type d'unités lexicales sur le même axe paradigmatique sans apporter des informations complètes, et surtout précises, sur les niveaux de langue et les différentes situations de communication, les dictionnaires créent plus de problèmes qu'ils n'en résolvent.

L'analyse que nous venons de proposer permettrait d'émettre l'hypothèse que la syntaxe serait suffisante pour mettre en évidence la combinatoire syntaxique et lexicale de certaines unités lexicales et en inférer leur niveau de langue.

Mais nous constatons en même temps que notre cadre théorique ne peut pas non plus résoudre tous les problèmes. Un bon exemple est celui des adjectifs et des locutions. Par exemple, il faut se demander quelles peuvent être les propriétés syntaxiques en français ou en grec qui nous permettraient de parler d'un

19 Selon Gross et Kiefer (1995), il s'agit de paramètres événementiels qui caractérisent les prédicats ayant un emploi événementiel.

20 En français, le verbe *baver* ne porte pas de marque dans le PR, alors qu'il fait partie de la nomenclature du dictionnaire de l'argot de Colin (1994). En revanche, *bavasser* et *débiner* se trouvent dans la macrostructure du PR, et sont aussi enregistrés dans Colin. Ils portent la marque *fam.* Le verbe *bavasser* est caractérisé, en plus, par la marque *péj.* dans le PR (2008).

21 Selon le PR, *causer*² est proche de *cancaner*, l'action de parler trop, avec indiscrétion, malignité. Notons que le paradigme de ce type de verbes est assez long dans les deux langues. Citons comme exemples les verbes *caqueter*, *jaboter*, *discourir*, *laïusser*, *déblatérer*, *casser du sucre sur le dos de qqn*, etc. ou bien *κακαρίζω*, *παράλω*, *λέω λόγια του αέρα*, *γλωσσοκοπανώ*, *τρώω γλυστρίδα*. Ce sujet mériterait une plus ample analyse, ce que nous nous proposons de faire dans une future étude.

niveau différent dans le cas de *hypocrite* et *faux cul*. Ainsi, tous deux peuvent être combinés avec *espèce de*. De même, il n'existe pas de propriétés syntaxiques distinctives pour des expressions comme *τα βρήκα αγγούρια* « tomber sur un os », *κάτι είναι αγγούρι* « quelque chose est emmerdant » ou *στέκεται σαν αγγούρι* « raide comme un piquet ».

CONCLUSION

On constate que les dictionnaires bilingues dont les langues source et cible sont le grec et le français présentent plusieurs problèmes concernant le choix des marques pour exprimer les niveaux de langue : certains dictionnaires utilisent des termes plutôt génériques (par exemple *populaire* dans le dictionnaire de Rosgovas) et d'autres utilisent les termes que l'on trouve dans la tradition lexicographique (par exemple, *argot*, *familier*, *populaire* et *vulgaire*, dans les dictionnaires des éditions Kauffmann). Mais même dans ce cas-là, il faut bien souligner que le choix ne semble pas être tout à fait cohérent ou répondre aux attentes d'une analyse stricte : certains mots ne sont pas marqués tandis que d'autres le sont dans le dictionnaire bilingue mais pas dans le monolingue. Quoi qu'il en soit, ce qui est évident c'est que dans tous ces dictionnaires les auteurs n'explicitent jamais leurs choix théoriques.

Par ailleurs, nous pouvons presque comprendre les raisons pour lesquelles les dictionnaires ont toutes ces difficultés : en fait, il s'agit tout simplement d'une classification qui ne représente pas vraiment la réalité linguistique et les limites de ces niveaux sont imposées par des facteurs extralinguistiques et subjectifs, propres parfois aux rédacteurs du dictionnaire.

Nous avons donc envisagé la possibilité de considérer comme une solution la description des propriétés syntaxiques. L'application d'une telle approche nous permet de distinguer les différents registres de langue à l'aide du contexte et des propriétés distributionnelles mais elle crée, elle aussi, de nouveaux problèmes. Même si cette approche peut être efficace avec certains prédicats nominaux (comme ceux qui expriment une *dispute* ou un *sentiment*), elle pose les mêmes problèmes dans le cas des adjectifs ou de certaines locutions.

En tout cas ce qui paraît inévitable et même obligatoire, c'est la mention dans le dictionnaire des critères selon lesquels les rédacteurs utilisent les marques de niveaux de langue ainsi qu'une définition des marques utilisées.

RÉFÉRENCES BIBLIOGRAPHIQUES

- AL Bernard, 1981, « Principes d'organisation d'un dictionnaire bilingue », *Lexique*, 2, p. 159-165.
- CALVET Louis-Jean et MATHIEU Patrick (éds), 2003, *Marges linguistiques*, 6, Argots, « français populaires » et langues populaires.
- CHAURAND Jacques et MAZIÈRE Françoise (éds), 1990, *La Définition*, Paris, Larousse.
- CHERKAoui-MESSIN Kenza, 2005, *Dictionnaires bilingues français / langues de l'UE*, rapport CNL et DGLFLF
www.dglf.culture.gouv.fr/recherche/Langues_du_monde.pdf

- CORBIN Pierre et GASIGLIA Nathalie (dans ce volume), « Éléments pour un état de la description de la variété des usages lexicaux dans les dictionnaires français monolingues actuels ».
- CRUSE David A., 1986, *Lexical Semantics*, Cambridge, Cambridge University Press.
- DUBOIS Jean, 1981, « Dictionnaire et syntaxe », *Lexique*, 2, p. 85-88.
- DUBOIS Jean et DUBOIS Claude, 1971, « Le dictionnaire et le bilinguisme », *Introduction à la lexicographie : le dictionnaire*, Paris, Larousse, p. 34-38.
- FRANÇOIS-GEIGER Denise, 1991, « Panorama des argots contemporains », *Langue française*, 90, p. 5-9.
- FRANÇOIS-GEIGER Denise et GOUDAILLIER Jean-Pierre (éds), 1991, *Langue française*, 90, *Parlures argotiques*.
- GADET Françoise, 2003, *La Variation sociale en français*, Paris, Ophrys.
- GROSS Gaston, 1995, « Une sémantique nouvelle pour la traduction automatique : les classes d'objets », *La Tribune des Industries de la Langue et de l'Information électronique*, 17-18-19, p. 16-19.
- GROSS Gaston, 1998, « Pour une véritable fonction "synonymie" dans un traitement de texte », *Langages*, 131, p. 103-114.
- GROSS Gaston et KIEFER Ferenc, 1995, « La structure événementielle des substantifs », *Folia Linguistica*, 29 (1-2), p. 43-65.
- HARRIS Zelig Sabbetai, 1976, *Notes du cours de syntaxe*, Paris, Le Seuil.
- HAUSSMAN Franz Josef, REICHMANN Oskar, WIEGAND Herbert Ernst und ZGUSTA Ladislav (Hrsg.), 1989-1991, *Wörterbücher / Dictionaries / Dictionnaires. Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie*, 3 vol., Berlin u. New York, Walter de Gruyter.
- LANDHEER Ronald, 1981, « Ambiguïté et dictionnaire bilingue », *Lexique*, 2, p. 147-158.
- LE PESANT Denis et MATHIEU-COLAS Michel (éds), 1998, *Langages*, 131, *Les classes d'objets*.
- LE PESANT Denis et MATHIEU-COLAS Michel, 1998, « Introduction aux classes d'objets », *Langages*, 131, p. 6-33.
- LÉPINETTE Brigitte, 1997, « Le rôle de la syntaxe dans la lexicographie bilingue », dans H. Béjoint et Ph. Thoiron, *Les Dictionnaires bilingues*, Louvain-la-Neuve, Aupelf-Uref – Duculot, p. 53-69.
- MARELLO Carla, 1997, « Les différents types de dictionnaires bilingues », dans H. Béjoint et Ph. Thoiron, *Les Dictionnaires bilingues*, Louvain-la-Neuve, Aupelf-Uref – Duculot, p. 31-52.
- MEL'CUK Igor, CLAS André et POLGUÈRE Alain, 1995, *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve, Aupelf-Uref – Duculot.
- PAULIN Aurélia, 1997, « Analyse de la notion de non-standard dans les dictionnaires bilingues français / anglais », *Meta*, 42 (1), p. 55-67.
- PETROPOULOS Elias, 1971, *Kaliarda: An Etymological Dictionary of Greek Homosexual Slang*, Athens, Kedros (Nefeli, 1980).
- ROBERTS R. P., 1994, « Marques de registres dans les dictionnaires bilingues », dans L. Mercier et C. Verreault, *Les Marques lexicographiques en contexte*

québécois, Québec, Gouvernement du Québec, Office de la langue française, coll. Etudes, Recherches et Documentation, p. 191-201.

ROUAYRENC Catherine, 1997, *Les Gros Mots*, Paris, Puf.

SZABÓ Dávid, 2001, « Les registres non conventionnels dans le dictionnaire bilingue », dans Anne-Marie Löffler-Laurian (éd.), *Études de linguistique générale et contrastive. Hommage à Jean Perrot*, Paris, Centre de recherche sur les langues et les sociétés, p. 427-436.

SZENDE Thomas, 1997, « Problèmes d'équivalence dans les dictionnaires bilingues », dans H. Béjoint et Ph. Thoiron, *Les Dictionnaires bilingues*, Louvain-la-Neuve, Aupelf-Uref – Duculot, p. 111-126.

SZENDE Thomas, 2000, *Approches contrastives en lexicographie bilingue*, Paris, Champion.

DICTIONNAIRES CONSULTÉS

LUST Colette et PANTELODIMOS Dimitris, 1995, *Dictionnaire français-grec moderne*, Athènes, Librairie Kauffmann, 120 000 entrées, 1 380 pages.

LUST Colette et PANTELODIMOS Dimitris, 2004 (2^e éd.), *Dictionnaire français-grec moderne*, Athènes, Librairie Kauffmann, 140 000 entrées, 1 112 pages.

PANTELODIMOS Dimitris et KÄITERIS Konstantinos, 2002, *Dictionnaire grec-français*, Athènes, Librairie Kauffmann, 150 000 entrées, 1 507 pages.

ROSGOVAS Theodore, 2002, *Dictionnaire grec-français*, 100 000 entrées, 649 pages (du grec vers le français seulement).

ROSGOVAS Theodore, 2002, *Dictionnaire français-grec*, 100 000 entrées, 681 pages (du français vers le grec seulement).

ROBERT Jean-Pierre et MALAMAS-ROBERT Maria, 2002, *Dictionnaire français-grec moderne*, Athènes, Efstathiadis (et Paris, Ophrys), nombre d'entrées non précisé, 807 pages.

BÉCOS Dimitris, 2005, *L'Argot moderne sans peine avec des dialogues authentiques*, Athènes, Editions Eiffel, 208 pages.

COLIN Jean-Paul, MÉVEL Jean-Pierre et LECLÈRE Christian, 2002, *Dictionnaire de l'argot français et de ses origines*, Paris, Larousse, 903 pages.

Le Nouveau Petit Robert (millésime 2008), Paris, Dictionnaires Le Robert, 2009, 60 000 entrées, 2 837 pages.

LES MARQUES D'USAGE DANS UN DICTIONNAIRE ÉLECTRONIQUE

Vassiliki FOUFI

Tita KYRIACOPOULOU

Claude MARTINEAU

Anastasia YANNACOPOULOU

Université Aristote de Thessalonique

Institut Gaspard-Monge, Université de Marne-la-Vallée

Résumé — Nous présentons ici les marques d'usage et leur utilité dans un dictionnaire électronique en nous appuyant sur deux cas concrets : le Dictionnaire électronique du français et le Dictionnaire électronique du grec moderne. Le premier est élaboré depuis une trentaine d'années au Laboratoire d'automatique documentaire et linguistique (LADL) et le deuxième depuis une dizaine d'années au Laboratoire de traduction et de traitement automatique du langage. Le dictionnaire français comporte aujourd'hui 240 000 entrées et le grec 170 000. On parlera en particulier : (1) des marques lexicographiques utilisées dans ces deux dictionnaires ; (2) de l'utilité de ces marques dans le cadre d'un dictionnaire électronique ; (3) de la difficulté à coder ces marques. La marque lexicographique sera présentée selon le dictionnaire. Par exemple le dictionnaire grec contient la marque +sv (forme savante). Rappelons que le grec est une langue marquée par la diglossie, langue savante (*katharevoussa*) et langue démotique, bien que, depuis plus de vingt ans, la première soit la langue officielle. Quant au dictionnaire du français, il inclut les couches lexicographiques (couche1 = mots courants, couche2 = mots moins courants, couche3 = mots rares ou techniques). On parlera en particulier du rôle de ces marques à travers les applications informatiques.

Mots clés — dictionnaire électronique, marques d'usage, couches lexicographiques, applications informatiques.

Abstract — In this article, we examine labels and their usefulness in electronic dictionaries, our study being based on two concrete examples: the Electronic Dictionary of French and the Electronic Dictionary of Modern Greek. The first has been developed over the past thirty years at the Laboratoire d'automatique documentaire et linguistique (LADL) and the second over the last decade at the Laboratoire de traduction et de traitement automatique du langage. The French dictionary presently contains 240 000 entries and the Greek about 170 000. Three items in particular are discussed: (1) lexicographical labels used in the

two dictionaries; (2) the usefulness of these labels as part of an electronic dictionary; (3) the difficulty in coding such labels. The study is focused on two labels: the label +sv (learned form) in the Greek dictionary and the label of *frequency* as used in French dictionary such as *common* words, *less common* words, and *rare* or *technical* words. It is to be noted that Greek is characterized by diglossia, i.e., opposing scholarly language (*katharevousa*) and demotic language, the former having been the official language for more than twenty years. The role these labels play in computer-based applications and the complexity of applying them are discussed.

Keywords — electronic dictionary, usage labels, learned form label, frequency lexicographical layers, computer-based applications.

INTRODUCTION

Dans cet article, nous présentons le dictionnaire électronique du français (DELA) qui est élaboré au Laboratoire d'automatique documentaire et linguistique (LADL) de l'Université de Marne-la-Vallée et celui du grec moderne (DELA-G) qui est construit par les membres du Laboratoire de traduction et de traitement automatique du langage de l'Université Aristote de Thessalonique.

Par dictionnaire électronique, nous entendons un dictionnaire utilisable dans des applications informatiques et non pas un dictionnaire sous forme informatique (saisi par traitement de texte par exemple). En effet, notre but est l'élaboration de dictionnaires électroniques pour des applications comme les correcteurs orthographiques et grammaticaux, l'analyse et la traduction automatique, et d'autres applications relatives au traitement automatique des textes (résumés automatiques de textes, filtrage d'informations, recherche intelligente sur le web, etc.).

Après une brève présentation des deux dictionnaires, nous nous focaliserons sur les marques lexicographiques codifiées en français et en grec. Nous parlerons en particulier de la marque +sv (= forme savante) utilisée dans le dictionnaire du grec pour qualifier les formes fléchies de *katharevousa* (section 3.1). Le dictionnaire du français inclut les couches lexicographiques (couche1 = mots de sens évident pour tout locuteur natif, couche2 = vocabulaire d'usage moins courant, couche3 = mots rares ou techniques) qui concernent la compréhension des mots par les locuteurs et non pas leur fréquence dans les textes (section 3.2).

1. DICTIONNAIRES ÉLECTRONIQUES

La constitution d'un dictionnaire électronique ne peut être qu'une tâche collective et de longue haleine. En effet, la construction du DELA a nécessité depuis plus de vingt-cinq ans l'intervention d'une dizaine de chercheurs et le DELA-G d'une quinzaine de chercheurs depuis quinze ans. Les deux dictionnaires sont toujours en cours de réalisation.

Dans les textes, les mots sont fléchis. Ainsi, nous avons développé des outils qui génèrent automatiquement toutes les formes fléchies des mots afin qu'on puisse les utiliser dans les applications de traitement automatique des textes écrits. En ce qui concerne le français, le dictionnaire des mots simples (DELAS) comporte 123 000 entrées d'où sont générées 985 000 formes fléchies (DELA-F),

alors que le dictionnaire des mots composés (DELAC) comporte 126 000 entrées d'où sont générées 270 000 formes fléchies (DELACF).

En ce qui concerne les dictionnaires grecs, celui des mots simples (DELAS_G) contient 170 000 entrées d'où sont générées 1 800 000 formes fléchies (DELA_F_G) et celui des mots composés (DELAC_G) contient 28 000 entrées d'où sont générées 140 000 formes fléchies (DELACF_G).

Pour la génération automatique des formes fléchies, nous avons développé le programme de flexion *GenereFlexion* (Courtois 1990, Mrabti 2001, Martineau 2003). Le programme utilise deux fichiers : le fichier des formes canoniques avec les codes flexionnels et le fichier des terminaisons. Le programme génère automatiquement un troisième fichier qui contient toutes les formes fléchies :

Exemples grecs¹ :

βουνό, βουνό. N: Nns: Ans: Vns

βουνού, βουνό. N: Gns

βουνών, βουνό. N: Gnp

βουνά, βουνό. N: Nnp: Anp: Vnp

Exemples français² :

actifs, actif. A: mp

active, actif. A: fs

actives, actif. A: fp

Pour l'élaboration de nos dictionnaires, nous avons procédé au recensement des dictionnaires existants mais aussi au dépouillement de corpus journalistiques disponibles sur le web.

Bien évidemment, et dans un souci d'exhaustivité, les deux dictionnaires sont continuellement enrichis. Cependant une telle opération pose des problèmes et il est, par exemple, parfois difficile de faire des choix quant aux entrées d'un dictionnaire ; on hésite encore entre deux orthographes (*εταιρεία* ou *εταιρία* / société) ou deux déclinaisons (*σκέψης* ou *σκέψεως* / de la pensée). Et même lorsque l'orthographe ou la déclinaison d'un mot est standardisée dans les grammaires et les dictionnaires, elle n'est pas toujours adoptée par les utilisateurs. Par exemple en grec le génitif singulier au féminin de l'adjectif *άγιος* (saint) est *άγιας*, mais les statistiques sur la fréquence des formes dans les corpus montrent que la forme *αγίας* est utilisée à 95 %. La reconnaissance des mots d'un texte doit tenir compte des nombreuses variantes graphiques et grammaticales et même des « fautes ». Mais doit-on considérer comme faute la mauvaise orthographe d'une forme employée par 70 % des utilisateurs ? C'est une question à laquelle nous pouvons difficilement répondre. Par exemple le mot *εξόφληση* (solde) se rencontre très souvent (plus de 6 000 occurrences) écrit *εξώφληση*. Notre stratégie consiste à prendre en compte l'évolution de la langue, mais aussi les diverses applications. Ainsi pour le traitement d'un texte nous

1 Où *N* désigne la catégorie grammaticale nom, *Nns* = nominatif neutre singulier, *Ans* = accusatif neutre singulier, *Vns* = vocatif neutre singulier, *Gns* = génitif neutre singulier, *Gnp* = génitif neutre pluriel, *Nnp* = nominatif neutre pluriel, *Anp* = accusatif neutre pluriel, *Vnp* = vocatif neutre pluriel.

2 Où *A* désigne la catégorie grammaticale adjectif, *mp* = masculin pluriel, *fs* = féminin singulier, *fp* = féminin pluriel.

automate qui associe des informations supplémentaires à des séquences reconnues (Silberstein 1993). Nous pouvons observer (v. le schéma ci-dessus, les parties grisées renvoient à des sous-graphes) que les différentes expressions qui expriment le temps comme par exemple : *πέντε και δεκαπέντε πρώτα λεπτά* (cinq heures quinze), *πέντε η ώρα ακριβώς* (cinq heures pile), *πέντε το απόγευμα* (cinq heures de l'après-midi), *πέντε παρά τέταρτο* (cinq heures moins le quart), etc., recevront le trait +*HEURE*.

2. LES MARQUES LEXICOGRAPHIQUES

Les deux dictionnaires étant construits sur les mêmes bases linguistiques (en prenant bien sûr en compte les particularités de chaque langue) mais indépendamment l'un de l'autre, il est normal que les marques d'usage soient différentes. Elles seront explicitées dans les sections qui suivent.

Il est important d'insister ici sur l'intérêt des marques d'usage dans un dictionnaire électronique. Rappelons que ces dictionnaires sont très exhaustifs (par définition il n'y a pas les contraintes du nombre des pages, de l'utilisation pratique d'un manuel, etc.) et que les différentes applications ne nécessitent pas d'utiliser toujours des données aussi volumineuses. Il est donc souvent souhaitable, voire nécessaire, de pouvoir sélectionner une partie du dictionnaire en fonction des besoins. L'analyse lexicale est peut-être l'application qui exige des dictionnaires les plus exhaustifs possible car on ne peut pas prévoir ce qu'il y aura dans le texte à analyser. Or, les formes (mots et signes de ponctuation) identifiées isolément, c'est-à-dire hors contexte, sont ambiguës. Ces ambiguïtés sont gênantes pour l'analyse automatique des textes et la majorité d'entre elles peuvent être éliminées avec une exploration du contexte ou une exploitation des marques lexicographiques.

Pour résumer, nous pouvons donc dire que les marques d'usage sont utiles dans le cadre d'un dictionnaire électronique pour, d'une part, construire des dictionnaires « sur mesure » en fonction des applications et, d'autre part, aider à la résolution des ambiguïtés, trop fréquentes lorsque les dictionnaires sont exhaustifs. Prenons comme exemple le verbe grec *αγαπάω/αγαπώ* (aimer). À la 2^e personne du présent de l'indicatif la forme *αγαπάτε* sera reliée aux deux formes canoniques citées ci-dessus et la machine ne saura pas quelle forme choisir.

3. LA MARQUE D'USAGE +SV DU DICTIONNAIRE DU GREC

L'orthographe et la grammaire de la langue naturelle sont moins bien définies que celles des mots des langages informatiques. Divers problèmes, comme la mise au pluriel, l'usage des lettres capitales, ne sont pas résolus systématiquement. Le problème de l'orthographe est beaucoup plus présent dans une langue marquée par la diglossie, comme le grec : langue savante (*katharevoussa*) et langue démotique (Jordanidou 1985). En effet, depuis le remplacement de la langue savante, en 1976, par la démotique comme langue officielle, on hésite encore entre deux orthographes. En grec les variantes des mots (verbes, noms, adjectifs, etc.) sont nombreuses et elles n'affectent pas uniquement les formes canoniques mais aussi les formes fléchies.

Dans le dictionnaire grec, le trait +sv signale les formes savantes (Kindt & Yannacopoulou 2006). Par exemple les formes *πρέσβευς* et *δικαστούς* sont codées avec la marque +sv :

πρέσβεως,πρέσβυς.N+sv:Gms
δικαστοῦ,δικαστής.N+sv:Gms:Gfs

Il n'y a pas de trait distinctif pour les formes de démotique puisque par déduction toute forme non codée +sv est une forme en démotique.

Depuis trente ans la langue démotique est la langue officielle en Grèce et pourtant, lors du traitement du grec, nous constatons que des éléments de la *katharevousa* persistent dans le grec standard. Examinons le tableau 1 ci-après qui donne des statistiques sur un journal quotidien et sur des articles écrits par les mêmes journalistes entre 1997 et 2003 :

	1997		2000		2003	
	Démotique	Savante	Démotique	Savante	Démotique	Savante
Auteur1	56,8	43,2	67,7	32,3	75,5	24,5
Auteur2	62,5	37,5	76,5	23,5	79,0	21,0

1.- Statistiques Langue démotique / Langue savante (en %)

Nous pouvons observer que même si les formes de *katharevousa* diminuent, elles persistent quand même et cela à un pourcentage de 20 % environ.

La principale difficulté de cette marque est qu'elle a été codée de façon intuitive et donc arbitraire car nous n'avons trouvé que peu de critères formels pour la formaliser. Et pourtant quand des règles formelles existent nos lexicographes les appliquent. Ces règles, qui ne dépassent pas la trentaine, concernent par exemple :

- les terminaisons :
Démotique Savante
-ας -ος
-εψα -ευσα
- l'accentuation des noms et des adjectifs :
Démotique Savante
ἄγγελου *αγγέλου* (ange)

mais pas des verbes (*αγόρασαν* - *αγοράσανε* / ils ont acheté).

Cependant toutes les variantes ne sont pas affectées par la marque +sv puisqu'il existe des formes différentes qui :

- sont toutes de la langue démotique
αλβανή - *αλβανίδα* (albanaise)
Νοέμβριος - *Νοέμβρης* (novembre)
- sont toutes de la langue savante
κυβερνώσας - *κυβερνώσης* (gouverné-e)
- incluent des erreurs
τέσσερις - *τέσσερεις* (quatre)

4. COUCHES LEXICALES DU DICTIONNAIRE FRANÇAIS

Les dictionnaires usuels contiennent à la fois des mots « connus » et des mots « moins connus » et même des mots techniques. La délimitation de couches de vocabulaire représentant le lexique gradué (du plus au moins essentiel) d'une population ou d'un groupe social correspond à un outil très utile dans le traitement automatique des langues pour l'élimination des ambiguïtés par exemple.

Le dictionnaire français a été réparti en trois couches de façon à privilégier éventuellement une partie du lexique au lieu de faire des recherches dans le dictionnaire entier. Ces couches ont été établies (Bonan-Garrigues 1993) en fonction d'une part, de la probabilité supposée d'apparition des mots dans les textes et, d'autre part, de la compréhension des mots par les locuteurs.

Elles sont codées comme suit :

Couche lexicale	Code	Exemple
1 : mots usuels et vocabulaire général	+z1	<i>participer</i> ,.V3+z1
2 : mots moins couramment utilisés	+z2	<i>narcissiquement</i> ,.ADV+z2
3 : mots rares ou vocabulaire technique	+z3	<i>molybdénite</i> ,.N2I+z3

2.- Les couches lexicales

Grâce à ces couches le mot « il », codé en français comme nom (ancienne monnaie turque) et pronom, dans un texte courant, ne retient que l'information de la catégorie grammaticale pronom.

Pour le grec, l'étude de la répartition du lexique en plusieurs couches a commencé récemment. En attendant, les entrées du DELAF grec non encore traitées ont reçu le code +z1. Une telle répartition exige un travail de recherche très important ; il va s'achever dans un futur proche.

Il est important de noter que les marques d'usage présentées ci-dessus peuvent être aisément utilisées dans les automates (graphes) présentés brièvement dans la section 2.

La présentation en parallèle des dictionnaires électroniques grec et français construits dans les deux laboratoires cités ci-dessus illustre la possibilité de décrire des langues différentes suivant le même formalisme. À côté des DELAFs présentés ici, les DELACFs contiennent des mots et formes composés. Ces dictionnaires, incorporés au système *UNITEX* élaboré pour l'analyse automatique des textes, permettent d'explorer un grand nombre de documents écrits. Ainsi, les études lexicographiques et linguistiques peuvent être fondées sur des attestations précises.

CONCLUSION

Les bases de données que représentent les dictionnaires électroniques constituent des ensembles linguistiques fort utiles pour suivre l'évolution d'une langue. À l'heure de la mondialisation et de la diffusion croissante de publications sur Internet, le développement de tels dictionnaires est nécessaire dans divers domaines, non seulement informatique mais aussi pédagogique, sociologique, etc. Les marques d'usage sont un plus appréciable dans de tels dictionnaires. Elles permettent, entre autres, de lever des nombreuses ambiguïtés que

les systèmes d'analyse automatique engendrent inévitablement. Nous nous efforcerons de compléter nos dictionnaires, en particulier celui du grec, et de l'enrichir avec d'autres marques d'usage (*familier / soutenu, standard / dialectal*, etc.) à condition de trouver des critères formels et de limiter les critères intuitifs.

RÉFÉRENCES BIBLIOGRAPHIQUES

- BONAN-GARRIGUES Mylène, 1993, *Méthode de paramétrage des dictionnaires et grammaires électroniques, Application à des systèmes interactifs en langue naturelle*, thèse de doctorat de l'Université Paris-Diderot - Paris 7.
- COURTOIS Blandine, 1990, « Un système de dictionnaires électroniques pour les mots simples du français », *Langue française*, 87 : 11-22.
- JORDANIDOU Anna, 1985, *La Diglossie en Grèce : Étude d'un cas précis, le participe*, thèse de troisième cycle, Université Paris-Diderot - Paris 7.
- KINDT Bastien and YANNAKOPOULOU Anastasia, 2006, "Literary words-automatic recognition in a Modern Greek journalistic corpus", *Proceedings of the 7th International Conference on Greek Linguistics (CD-rom), 10-13 sept. 2005*, York University.
- MARTINEAU Claude, 2005, *Outils multilingues de génération de formes fléchies et dérivées*, Rapport technique Transweb 2, Université de Marne-la-Vallée - Paris 13.
- MRABTI Safia, 2001, *Flexion automatique des mots composés en grec*, mémoire de DEA, Université de Marne-la-Vallée - Paris 13.
- SILBERZTEIN Max, 1993, *Dictionnaires électroniques et analyse automatique de textes. Le système Intex*, Paris, Masson.
- VOYATZI Stavroula, 2006, *Description morpho-syntaxique et sémantique des ad-verbales figés de phrase en vue d'un système d'analyse automatique des textes grecs*, thèse de doctorat de l'Université de Marne-la-Vallée - Paris 13.

LES AUTEURS

Anna ANASTASSIADIS-SYMÉONIDIS est professeur au Département de Linguistique de l'Université Aristote de Thessalonique. Elle est auteur de sept livres et quatre-vingts articles concernant l'analyse du grec moderne : néologie, emprunt, morphologie dérivationnelle (suffixation, composition), figement, dictionnaires papier et électroniques, lexicographie, métalexigraphie. Dans le cadre de Pythagoras II, elle travaille sur un dictionnaire trilingue de cinq cents proverbes (grec-français-anglais), sur un dictionnaire étymologique des toponymes arvanites et sur la constitution d'une banque de données de néologismes grecs. Dans le cadre d'EPEAEK, elle travaille sur l'enseignement du grec langue seconde aux élèves de collèges grecs.

Jean-Claude ANSCOMBRE est directeur de recherches au CNRS. Il travaille dans la ligne du paradigme autonomiste du *Cours de linguistique générale* de Saussure. Dans les années 1970-1980, il fonde avec Oswald Ducrot une conception originale de la pragmatique, dite « pragmatique intégrée », inscrite dans le système de la langue et dans son fonctionnement, puis une approche sémantique originale : la « linguistique de propriétés », soucieuse de démonstration formelle, impliquant syntaxe, lexique, voire phonologie et métrique. Il s'est plus récemment consacrée à une « théorie des stéréotypes » destinée à rendre compte de l'identité des lexèmes et de leur fonctionnement en discours.

Dorothee AQUINO-WEBER est assistante de recherche au Centre de dialectologie et d'étude du français régional de l'Université de Neuchâtel. Elle travaille actuellement à sa thèse de doctorat en sociolinguistique historique française sous la direction du Pr. Andres Kristol : « *Les Mémoires d'un forban philosophe* (1829) : Témoignage littéraire des représentations linguistiques du début du XIX^e siècle ». Sa participation à des projets de recherche incluent celui de la Base de données lexicographiques panfrancophone (BDLPSuisse) et l'élaboration des index du second volume des *Documents linguistiques de la Suisse romande* sous la direction de Jean-Daniel Morerod, Rémy Scheurer et Wulf Müller.

Haïfa BEN MAHFOUDH-HUBERT est doctorante en sciences du langage à l'Université de Limoges. Elle travaille sur la description sémantique des mots et l'apport des proverbes dans le cadre de la réalisation d'un dictionnaire plurilingue des idéologies lexicalisées.

Nathalie GASIGLIA est maître de conférences habilitée en linguistique française et informatique, chercheuse en métalexigraphie et en traitement automatique

des corpus, auteur de diverses publications dans ces domaines, membre de l'UMR 8163 (Savoirs, Textes, Langage) du CNRS, coresponsable du Master « Lexicographie, Terminographie et Traitement Automatique des Corpus » à l'université Charles de Gaulle - Lille 3 (France).

Erzsébet CHMELIK est docteur de l'Université de Limoges. Elle travaille sur la construction de tests linguistiques basés sur la description des connecteurs en français et en hongrois. Elle est l'auteur d'une dizaine de publications dans le domaine de la sémantique, de l'analyse de discours et de la pragmatique.

Pierre CORBIN est professeur de linguistique française, chercheur en métalexigraphie et auteur de nombreuses publications dans ce domaine, membre de l'UMR 8163 (Savoirs, Textes, Langage) du CNRS, cofondateur et directeur de la revue *Lexique* (Presses Universitaires du Septentrion), coéditeur scientifique de la collection *Lexicographica Series Maior* (Walter de Gruyter), cofondateur et coresponsable du Master « Lexicographie, Terminographie et Traitement Automatique des Corpus » à l'université Charles de Gaulle - Lille 3 (France).

Christophe CUSIMANO est docteur de l'Université de Metz avec une thèse qui a pour titre « Polysémie et noms de sentiments. Essai de linguistique générale » (*La polysémie : Essai de sémantique générale*, Paris, L'Harmattan, 2009). Il est aussi l'auteur de contributions articulant sémantique et pragmatique, parues et à paraître dans les actes de nombreux colloques internationaux.

Jean-Nicolas DE SURMONT est chercheur autonome et docteur en sciences du langage de Paris-X Nanterre. Il est diplômé de l'Université Laval à Québec, de l'Université Charles-De-Gaulle-Lille III, de l'Université Pompeu-Fabra, de l'Université Paris X-Nanterre et a complété son post-doctorat à Louvain-la-Neuve. Ses travaux touchent à la sociolinguistique, la métalexigraphie, l'histoire culturelle et politique du Québec, la théorie de la poésie vocale. Il vient de publier chez ENS éditions son cinquième ouvrage intitulé « Vers une théorie des objets-chansons » qui entend rénover la théorie de la poésie orale tel qu'élaborée par Paul Zumthor. Il a aussi publié de nombreux articles sur le français en Belgique et au Québec. Il est l'auteur de plusieurs centaines de recensions de disques et d'ouvrages notamment pour les magazines spécialisés comme *Cap-aux-Diamants* et *Ecouter-voir* mais aussi pour des revues savantes. Ses travaux sont publiés dans 22 pays et dans 5 domaines différents. Conférencier et enseignant éclectique, il a enseigné tout au long de son parcours à des enfants âgés de 4 à 102 ans. Il a aussi été chargé d'enseignement à l'Université Nancy I et à l'Université de Metz et a été collaborateur scientifique à l'Université Libre de Bruxelles en 2007.

Lola-Laurence DEVOLDER a soutenu sa thèse, sous la direction d'Aziza Boucherit, à l'Université Paris Descartes en 2007. Son travail met au jour les critères qui font des structures linguistiques familières des réalisations perçues comme non conventionnelles, à partir d'une analyse énonciative et pragmatique mais également strictement linguistique. Ses recherches s'orientent vers la morphologie lexicale, l'analyse syntaxico-sémantique et la sociolinguistique, la compréhension des faits et effets de langue en discours permettant de rendre compte plus précisément des emplois et des évaluations qu'en font les locuteurs.

Mireille ELCHACAR est en thèse de doctorat à l'Université de Sherbrooke au Québec et l'Université de Cergy-Pontoise en France, sous la codirection de Louis Mercier et Jean Pruvost. Sa recherche porte sur les problématiques liées au traitement du vocabulaire politique dans un dictionnaire général destiné aux francophones du Québec. Elle travaille également au projet FRANQUS, qui vise à produire un dictionnaire général du français à l'intention des francophones du Québec. Elle y rédige entre autres le vocabulaire sociopolitique.

Vassiliki FOUFFI prépare son doctorat à l'Université Aristote de Thessalonique. De 2005 à 2006, elle a travaillé sur le projet « Développement d'un correcteur orthographique du grec moderne » en collaboration avec l'Université de Chypre. Membre du Laboratoire de traduction et de traitement automatique du langage (Département de langue et de littérature françaises) elle s'occupe de la construction des dictionnaires électroniques des mots simples et composés du grec moderne.

Nathalie GASIGLIA est maître de conférences habilitée en linguistique française et informatique, chercheuse en métalexigraphie et en traitement automatique des corpus, auteur de diverses publications dans ces domaines, membre de l'UMR 8163 (Savoirs, Textes, Langage) du CNRS, coresponsable du Master « Lexicographie, Terminographie et Traitement Automatique des Corpus » à l'université Charles de Gaulle - Lille 3 (France).

Tita KYRIACOPOULOU a soutenu son habilitation en 2003. Professeur à l'Université Paris-Est Marne-la-Vallée, elle a construit un dictionnaire électronique des verbes grecs et développé le logiciel LEXP pour la création et la gestion des bases de données terminologiques.

Efi LAMPROU est maître de conférences au département d'études françaises et de langues vivantes de l'Université de Chypre. Sa thèse de doctorat analyse les propriétés syntactico-sémantiques des prédicats nominaux de « réciprocité ». Ses recherches portent sur l'analyse des prédicats nominaux du point de vue syntaxique et sémantique, l'acquisition du vocabulaire de ce type en FLE et la traduction.

Claude MARTINEAU est docteur de l'Université de Marne-la-Vallée – Paris XIII et membre de l'Institut Gaspard Monge où il a développé le logiciel *Gr_Flex* pour la génération de dictionnaires électroniques des formes fléchies du grec moderne. Il travaille sur la création et le traitement de ressources lexicales françaises et grecques, sur la construction de dictionnaires électroniques et la génération des formes fléchies.

Camille MARTINEZ est chargé de cours à l'Université de Cergy-Pontoise et membre de l'association ÉROFA (Études pour une rationalisation de l'orthographe française d'aujourd'hui). Il est l'auteur d'une thèse, soutenue en 2009 et à paraître chez Honoré Champion, sur l'évolution de l'orthographe dans les dictionnaires de la fin du XX^e et du début du XXI^e siècle. Il a publié différents articles dans les *Cahiers de lexicologie* et dans les *Études de linguistique appliquée*. Il est secrétaire de rédaction de la *Revue française de linguistique appliquée* (RFLA) et rédacteur associé des *Études de linguistique appliquée* (Klincksieck).

Louis MERCIER a commencé sa carrière de chercheur au Trésor de la langue française au Québec (Université Laval), où il a codirigé le projet *d'Index lexicologique québécois* (1986), puis participé à la rédaction du *Dictionnaire historique du français québécois* (1998). Professeur titulaire au Département des lettres et communications de l'Université de Sherbrooke, il enseigne la linguistique française depuis 1994 (histoire et variation du français, lexicographie, français québécois...) et dirige le Centre d'analyse et de traitement informatique du français québécois (CATIFQ). Il est également membre du Laboratoire de lexicologie et lexicographie québécoises (LexiQué). Ses publications récentes portent sur l'histoire des dictionnaires québécois et sur le traitement lexicographique des noms d'espèces naturelles.

Christel NISSILLE occupe un poste « jeune chercheur » financé par le FNS au *Französisches Etymologisches Wörterbuch* (FEW), ATILF (Nancy Université-CNRS) et est chargée de cours à l'Université de Bâle (Institut für Französische Sprach- und Literaturwissenschaft). Elle écrit actuellement sa thèse de doctorat intitulée « “Grammaire floue” et enseignement du français en Angleterre au XV^e siècle : les leçons du manuscrit Oxford Magdalen 188 » (sous la direction du Pr. Andres Kristol, Neuchâtel, et du Pr. Jean-Paul Chauveau, Nancy). Les projets lexicographiques auxquels elle participe s'inscrivent dans le cadre du volet suisse de la Base de données lexicographiques panfrancophone (BDLPSuisse) et de la refonte partielle des premiers tomes du FEW.

Pavlos PAVLOU a étudié les langues et littératures étrangères à l'Université de Vienne en Autriche, s'est diplômé en anglais langue étrangère et en allemand à l'Université d'Illinois à Carbondale et a soutenu sa thèse de doctorat (PhD) à l'Université de Georgetown (Washington DC). Il a été maître de conférences au département d'études anglaises de l'Université de Chypre jusqu'à son décès inopiné en août 2010. Ses recherches ont porté sur l'enseignement des langues étrangères et les tests linguistiques, la linguistique appliquée et la sociolinguistique. Il était représentant de la République de Chypre à l'Unesco pour la Convention sur la Sauvegarde du patrimoine culturel immatériel.

Bert PEETERS (PhD, Australian National University, 1989) enseigne le français et la linguistique française en Australie (University of Tasmania, 1989-2006 ; Macquarie University, 2007-). Ses publications incluent des livres, des articles et des comptes rendus dans des domaines aussi divers que la linguistique générale, la sémantique, la syntaxe et la sociolinguistique françaises, la métalangue sémantique naturelle et la communication interculturelle franco-australienne. Bert Peeters prépare en ce moment une monographie et un ouvrage collectif sur le rapport entre langue et valeurs culturelles et sur la lumière que peut jeter sur ce rapport la métalangue sémantique naturelle.

Alena PODHORNÁ-POLICKÁ est maître de conférences à l'Université Masaryk de Brno (République tchèque). Sa thèse en cotutelle (Université Paris Descartes / Université Masaryk) porte sur les universaux argotiques, plus spécifiquement sur les pratiques et les représentations des jeunes Français et Tchèques quant à leur argot groupal, régional ou générationnel. Elle travaille sur l'analyse linguistique des films de banlieue, sur les intensificateurs dans l'argot des jeunes et sur la traduction des chansons de rap en tchèque.

Pierre-Yves RACCAH, chercheur en linguistique au CNRS, a aussi enseigné la sémantique dans plusieurs universités françaises et étrangères. Ancien Président de l'Association pour la Recherche Cognitive, il est l'auteur ou le coauteur d'une centaine de publications dans des domaines liés à la sémantique, aux sciences de la cognition et à l'épistémologie de la linguistique.

Jackie SCHÖN est maître de conférences honoraire au département des sciences du langage et membre du laboratoire de psycholinguistique Jacques-Lordat de l'Université de Toulouse le Mirail, de la Société Internationale de Linguistique Fonctionnelle (SILF) et de l'International Association of Dialogue Analysis (IADA, Bologna & Munster). Ses domaines de recherche comprennent, entre autres, la familiarité langagière, l'interface Linguistique et Psychanalyse ainsi que celle de Sexe et Genre.

Freiderikos VALETOPOULOS est maître de conférences à l'Université de Poitiers (Linguistique française et Français Langue Etrangère). Après une thèse de doctorat consacrée à la classification sémantique des adjectifs prédicatifs en grec et en français, fondée sur leurs propriétés syntaxiques, il travaille actuellement sur les prédicats de sentiment et de perception. Ses recherches portent pour l'essentiel sur l'interface de la syntaxe et de la sémantique et l'acquisition du français langue étrangère par les apprenants hellénophones.

Anastasia YANNACOPOULOU est docteur de l'Université de Marne-la-Vallée – Paris XIII. Elle occupe un poste de linguiste informaticienne à Systran. Elle s'intéresse en particulier au traitement automatique des langues naturelles, à la construction de dictionnaires électroniques monolingues et multilingues et à la gestion des données linguistiques dans les dictionnaires électroniques.

TABLE DES MATIÈRES

Introduction

par Fabienne Hélène BAIDER, Efi LAMPROU et Monique MONVILLE-BURSTON	7
--	---

PREMIÈRE PARTIE ASPECT SÉMANTIQUE

I. TYPOLOGIE

1. Éléments pour un état de la description de la variété des usages lexicaux dans les dictionnaires français monolingues (1980-2008) par Pierre CORBIN et Nathalie GASIGLIA	17
2. Le poids des contraintes dictionnairiques sur l'évolution des marqueurs dans les <i>Petit Larousse</i> (1997-2007) par Camille MARTINEZ	39
3. Les marques d'usage dans le dictionnaire du grec standard par Anna ANASTASSIADIS-SYMÉONIDIS	51

II. ÉQUIVALENCE

4. Classification des formes sentencieuses et traduction d'énoncés parémiques par Jean-Claude ANSCOMBRE	67
5. Les faux amis, une question de degré : l'apport de la métalangue sémantique naturelle par Bert PEETERS	87
6. La traduction en lexicographie bilingue : des marques de catégorisation implicites ? par Christophe CUSIMANO	111
7. Les points de vue dans le lexique et dans le dictionnaire par Haïfa BEN MAHFOUDH-HUBERT, Erzsébet CHMELIK et Pierre-Yves RACCAH	117

DEUXIÈME PARTIE VALEURS D'EMPLOI

I. MARQUE DIATOPIQUE

8. Prolégomènes à une étude des marques d'usage dans la lexicographie différentielle : le cas du français de Suisse romande
par Dorothee AQUINO-WEBER et Christel NISSILLE 139
9. Mise en relation des particularismes québécois et hexagonaux dans les dictionnaires usuels du français : vocabulaire politique et limites du marquage topolectal
par Mireille ELCHACAR et Louis MERCIER 153
10. Contourner la marque diatopique : vitalité des emplois et présence des topolectismes dans les dictionnaires de variétés de français
par Jean Nicolas DE SURMONT 171
11. Simply too many things to label! Dealing with the lexicographical needs of bidialectal speech communities
by Pavlos PAVLOU 183

II. MARQUE DIASTRATIQUE

12. *Arg., Pop., Fam.* : Trois œufs du même panier ?
par Lola-Laurence DEVOLDER 197
13. L'expressivité et la marque lexicographique : étude comparative franco-tchèque d'un corpus du lexique non standard. Les marques *fam., pop., arg. vs expressivité* en lexicographies française et tchèque
par Alena PODHORNÁ-POLICKÁ 209
14. Du mécanisme caché sous la marque lexicographique *fam.* : le cas des noms
par Jackie SCHÖN 227
15. Les marques de *niveaux de langue* et la lexicographie bilingue français-grec
par Efi LAMPROU et Freiderikos VALETOPOULOS 235
16. Les marques d'usage dans un dictionnaire électronique
par Vassiliki FOUFI, Tita KYRIACOPOULOU, Claude MARTINEAU
et Anastasia YANNAKOPOULOU 253

Les auteurs 261